

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 19, 2002, 16:32:27 ; Search time 53.19 Seconds
(without alignments)
682.856 Million cell updates/sec

Title: US-09-771-961-2

Sequence: 1 MCSSTGCLLEPLDDDLN.....WIOGHGWEKTLGISHEVD 327

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Lasting first 45 summaries

Database :

- 1: /Genesec.032802:*
- 2: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1980.DAT:*
- 3: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1981.DAT:*
- 4: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1982.DAT:*
- 5: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1983.DAT:*
- 6: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1984.DAT:*
- 7: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1985.DAT:*
- 8: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1986.DAT:*
- 9: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1987.DAT:*
- 10: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1988.DAT:*
- 11: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1989.DAT:*
- 12: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1990.DAT:*
- 13: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1991.DAT:*
- 14: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1992.DAT:*
- 15: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1993.DAT:*
- 16: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1994.DAT:*
- 17: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1995.DAT:*
- 18: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1996.DAT:*
- 19: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1997.DAT:*
- 20: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1998.DAT:*
- 21: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA1999.DAT:*
- 22: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2000.DAT:*
- 23: /SIDSL/gcgdata/hold-genesec/genesecp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	1709	100.0	327	AAE07040 Human BCL-X-like p
2	1709	100.0	327	AAE07040 Human BCL-X-like p
3	1189.5	69.6	252	AAE07041 Human BCL-X-like p
4	1189.5	69.6	252	AAE07041 Human BCL-X-like p
5	929.5	54.4	328	AAE07041 Human BCL-X-like p
6	691.5	40.5	151	AAE07041 Human BCL-X-like p
7	686	40.1	129	AAE07041 Human BCL-X-like p
8	686	40.1	129	AAE07041 Human BCL-X-like p
9	686	40.1	129	AAE07041 Human BCL-X-like p
10	686	40.1	129	AAE07041 Human BCL-X-like p
11	686	40.1	129	AAE07041 Human BCL-X-like p

12	686	40.1	129	22	AAE07040	Human BCL-X-like p
13	686	40.1	129	22	AAE07040	Human BCL-X-like p
14	686	40.1	129	22	AAE07040	Human BCL-X-like p
15	686	40.1	129	22	AAE07040	Human BCL-X-like p
16	686	40.1	129	22	AAE07040	Human BCL-X-like p
17	686	40.1	129	22	AAE07040	Human BCL-X-like p
18	686	40.1	129	22	AAE07040	Human BCL-X-like p
19	686	40.1	129	22	AAE07040	Human BCL-X-like p
20	686	40.1	129	22	AAE07040	Human BCL-X-like p
21	686	40.1	129	22	AAE07040	Human BCL-X-like p
22	686	40.1	129	22	AAE07040	Human BCL-X-like p
23	686	40.1	129	22	AAE07040	Human BCL-X-like p
24	686	40.1	129	22	AAE07040	Human BCL-X-like p
25	686	40.1	129	22	AAE07040	Human BCL-X-like p
26	686	40.1	129	22	AAE07040	Human BCL-X-like p
27	686	40.1	129	22	AAE07040	Human BCL-X-like p
28	686	40.1	129	22	AAE07040	Human BCL-X-like p
29	686	40.1	129	22	AAE07040	Human BCL-X-like p
30	686	40.1	129	22	AAE07040	Human BCL-X-like p
31	686	40.1	129	22	AAE07040	Human BCL-X-like p
32	686	40.1	129	22	AAE07040	Human BCL-X-like p
33	686	40.1	129	22	AAE07040	Human BCL-X-like p
34	686	40.1	129	22	AAE07040	Human BCL-X-like p
35	686	40.1	129	22	AAE07040	Human BCL-X-like p
36	686	40.1	129	22	AAE07040	Human BCL-X-like p
37	686	40.1	129	22	AAE07040	Human BCL-X-like p
38	686	40.1	129	22	AAE07040	Human BCL-X-like p
39	686	40.1	129	22	AAE07040	Human BCL-X-like p
40	686	40.1	129	22	AAE07040	Human BCL-X-like p
41	686	40.1	129	22	AAE07040	Human BCL-X-like p
42	686	40.1	129	22	AAE07040	Human BCL-X-like p
43	686	40.1	129	22	AAE07040	Human BCL-X-like p
44	686	40.1	129	22	AAE07040	Human BCL-X-like p
45	686	40.1	129	22	AAE07040	Human BCL-X-like p

ALIGNMENTS

RESULT 1
ID AAE07040 standard; Protein: 327 AA.
AC AAE07040;
XX 23-OCT-2001 (first entry)
XX Human BCL-X-like protein #1.
XX Human BCL-X-like protein; therapy: physiological disorder.
XX Homo sapiens.
XX WO200157213-A2.
XX 09-AUG-2001
XX 31-JAN-2001; 2001WO-US03446.
XX 04-FEB-2000; 2000US-0180412.
XX (LEXI-) LEXICON GENETICS INC.
XX Donoho G, Hilbun E, Turner CA, Friedrich G, Abuin A, Zambrowicz B;
XX Sands AT;
XX WPI: 2001-488882/53.
XX N-PSDB: AAD13235.
XX Novel isolated human BCL-X-like polynucleotide, useful in therapeutic,
XX diagnostic and pharmacogenic applications
XX Claim 4; Page 30-31; 33pp; English.

XX The present sequence is human BCL-X-like protein.
 CC The BCL-X-like polynucleotides are useful in therapeutic, diagnostic
 CC and pharmacogenetic applications. They are useful for screening drugs
 CC effective in the treatment of symptomatic or phenotypic manifestations
 CC perturbing the normal function of protein in the body and also for
 CC treating physiological disorders and diseases. The BCL-X-like
 CC polynucleotides are useful in conjunction with polymerase chain
 CC reaction to screen libraries, isolate clones, to prepare cloning
 CC and sequencing templates and as hybridisation probes for assessing
 CC gene expression patterns.
 XX

Sequence 327 AA:

Query Match
 Best Local Similarity 100.0%; Score 1709; DB 22; Length 327;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MCSTSGCDLEIPDDDDLTIEFKIAYTRHHVFKSTPALFSPKILRTSRISQRLGN 60
 Db 1 mcsstsgcdleipdddddltiefkilaytrhhvfksptalfspkllrtsrissqrlgn 60
 QY 61 CSANESWTEVSWPCRNSSQSEKAINLGKKSSWKAFFGVYKEDSOSTPAKVSAGQRTL 120
 Db 61 csaneswtevswpornsgssekaingkksswkaffgvvekedsgstpakvsagqrtl 120
 QY 121 EYDSSHQSQMSRCLSNVEOCLHEAVDPKYSTANRVAELIVSWPPQATQAGGFSKEI 180
 Db 121 eydshsqmsrcslsnveocleheavdpkystanrvaelivswppqatqagfskei 180
 QY 181 FYTEGSLFQLOGHVPYASSSKKDEBQILAKIVLELKYSGGDLERKLRKDKALMGHPDQ 240
 Db 181 fteglstfqlghvpyassskkdeeqllakivelkysgdlerkllrkdkalmghpdq 240
 QY 241 LSYVFKRTTDDVLMGVDPGRGESEVKAQGFKAALVIDYAKTALTDHNPNNRVLGFGTKY 300
 Db 241 lsvyfkrttdvmlgvdpgrgesevkaagfkfaalvidyaktaltdhnpnnrvlgfgtky 300
 QY 301 LKENSPWIOQHGWGWEKILGISHEEVD 327
 Db 301 lkenfswiqhgwgeklilgisheevd 327

RESULT 2

AAE07041
 ID AAE07041 standard; Protein; 327 AA.
 AC AAE07041;

QY 307-SEP-2001 (first entry)

Db Human BCL-G1 polypeptide.

XX BCL-G; cancer; cancer therapy; oncogene; apoptosis; BCL-G1; cytostatic;
 XX antiapoptotic; chromosome 12p12.3; human.

OS Homo sapiens.

PN WO20014282-A2.

PD 21-JUN-2001.

PF 13-DEC-2000; 2000WO-US33793.

PR 14-DEC-1999; 99US-0461641.

PA (BURN-) BURHAM INST.

PI Reed JC, Godzik A;

DR WPI; 2001-398125/42.

DR N-PSDB; AAH22582.

XX Novel polynucleotide encoding BCL-G polypeptide, useful for modulating
 PT apoptosis, and for diagnosing and treating cancer -
 PS Claim 14; Fig 2; 117pp; English.

XX The invention relates to BCL-G polypeptides and nucleic acids encoding
 CC them. The BCL-G polypeptides can be expressed by standard recombinant
 CC methodology. BCL-G oligonucleotides (or its anti-sense strand) and BCL-G
 CC specific antibodies are useful for diagnosing cancer, monitoring cancer
 CC therapy or assessing prognosis of patients with cancer. The BCL-G
 CC polypeptides are useful for modulating the activity of an oncogenic
 CC level of apoptosis mediated by the BCL-G polypeptide. A therapeutic
 CC composition comprising the BCL-G polypeptide, polynucleotide or antibody
 CC is useful for treating a pathology characterized by abnormal cell
 CC proliferation especially cancer. The present sequence represents a
 XX human BCL-G1 polypeptide.

Sequence 327 AA:

Query Match
 Best Local Similarity 100.0%; Score 1709; DB 22; Length 327;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MCSTSGCDLEIPDDDDLTIEFKIAYTRHHVFKSTPALFSPKILRTSRISQRLGN 60
 Db 1 mcsstsgcdleipdddddltiefkilaytrhhvfksptalfspkllrtsrissqrlgn 60
 QY 61 CSANESWTEVSWPCRNSSQSEKAINLGKKSSWKAFFGVYKEDSOSTPAKVSAGQRTL 120
 Db 61 csaneswtevswpornsgssekaingkksswkaffgvvekedsgstpakvsagqrtl 120
 QY 121 EYDSSHQSQMSRCLSNVEOCLHEAVDPKYSTANRVAELIVSWPPQATQAGGFSKEI 180
 Db 121 eydshsqmsrcslsnveocleheavdpkystanrvaelivswppqatqagfskei 180
 QY 181 FYTEGSLFQLOGHVPYASSSKKDEBQILAKIVLELKYSGGDLERKLRKDKALMGHPDQ 240
 Db 181 fteglstfqlghvpyassskkdeeqllakivelkysgdlerkllrkdkalmghpdq 240
 QY 241 LSYVFKRTTDDVLMGVDPGRGESEVKAQGFKAALVIDYAKTALTDHNPNNRVLGFGTKY 300
 Db 241 lsvyfkrttdvmlgvdpgrgesevkaagfkfaalvidyaktaltdhnpnnrvlgfgtky 300
 QY 301 LKENSPWIOQHGWGWEKILGISHEEVD 327
 Db 301 lkenfswiqhgwgeklilgisheevd 327

RESULT 3

AAE07041
 ID AAE07041 standard; Protein; 252 AA.

AC AAE07041;

QY 23-OCT-2001 (first entry)

Db Human BCL-X-like protein #2.

XX Human; BCL-X-like protein; therapy; physiological disorder.

OS Homo sapiens.

PN WO200157213-A2.

PD 09-AUG-2001.

PF 31-JAN-2001; 2001WO-US03446.

PR 04-FEB-2000; 2000US-0180412.

Thu Jun 20 13:14:55 2002

us-09-771-961-2.rag

PA (LEXI-) LEXICON GENETICS INC.
 XX Donoho G, Hilbun E, Turner CA, Friedrich G, Abuin A, Zambrowicz B;
 PI Sands AT;
 XX WPI: 2001-488882/53.
 DR N-PSDB: AAD13236.
 PT Novel isolated human BCL-X-like polynucleotide, useful in therapeutic,
 PT diagnostic and pharmacogenic applications
 PS Claim 5; page 32; 33pp; English.
 CC The present sequence is human BCL-X-like protein.
 CC The BCL-X-like polynucleotides are useful in therapeutic, diagnostic
 CC and pharmacogenic applications. They are useful for screening drugs
 CC effective in the treatment of symptomatic or phenotypic manifestations
 CC perturbing the normal function of protein in the body and also for
 CC treating physiological disorders and diseases. The BCL-X-like
 CC polynucleotides are useful in conjunction with polymerase chain
 CC reaction to screen libraries, isolate clones, to prepare cloning
 CC and sequencing templates and as hybridisation probes for assessing
 CC gene expression patterns.
 CC
 SQ Sequence 252 AA:

Query Match 69.6%; Score 1189.5; DB 22; Length 252;
 Best Local Similarity 93.5%; Pred. No. 6.4e-113;
 Matches 232; Conservative 2; Mismatches 7; Indels 7; Gaps 1;

QY 1 MCSTSGCDLEFIPDDDDNTIEFKILAYTRHHVFKSTPAFPKILRTSLSGRGLGN 60
 1 mcsstsgcdleipdddddntiefkilaaytrhhvfkstpalispkllrtslsgrglgn 60
 DB 61 CSANESWTEVSWPCRNSSQSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKYSAQGRFL 120
 61 csaneswtevwpcrnssqsekaingkksswkaffgvvekedsgstpkysaaggrfl 120
 QY 121 EYQDSHSQOORCLSNVBOCLEHEAVDPKVISTANRYAEIYVSWPPQATQAGFKSKEL 180
 121 eyqdsnsgwscrlsnvboacleheavdpkvistanyaeiyswppqatqagfkskel 180
 DB 121 EYQDSHSQOORCLSNVBOCLEHEAVDPKVISTANRYAEIYVSWPPQATQAGFKSKEL 180
 121 eyqdsnsgwscrlsnvboacleheavdpkvistanyaeiyswppqatqagfkskel 180
 QY 181 FYTEGLSFQLOGHVPYASSSKKDEEQLIAKIVELLKYSQGLERKRLKK-----DKAL 233
 181 fyteglsfqloghvpayassskkdeeqllakivellkysgqlerkdtafipivdtsi 240
 DB 181 fyteglsfqloghvpayassskkdeeqllakivellkysgqlerkdtafipivdtsi 240
 181 fyteglsfqloghvpayassskkdeeqllakivellkysgqlerkdtafipivdtsi 240
 QY 234 MGHFODGL 241
 234 mghfodgl 241
 DB 241 qgfpgdgl 248

RESULT 4
 ID AAB85167 standard; Protein: 252 AA.
 XX AAB85167;
 AC AAB85167;
 DT 07-SEP-2001 (first entry)
 XX Human Bcl-Gs polypeptide.
 DE Human Bcl-Gs polypeptide.
 XX Bcl-G; cancer; cancer therapy; oncogene; apoptosis; Bcl-Gs; cytostatic;
 KW antiapoptotic; chromosome 12p12.3; human.
 KM
 XX Homo sapiens.
 OS
 XX MO200144282-A2.
 PN
 XX 21-JUN-2001.
 PD
 XX 13-DEC-2000; 2000WO-US33793.
 PF
 XX

PR 14-DEC-1999; 99US-0461641.
 XX (BURN-) BURHAM INST.
 PA Reed JC, Godzik A;
 XX WPI: 2001-398125/42.
 DR N-PSDB: AAH22583.
 PT Novel polynucleotide encoding Bcl-G polypeptide, useful for modulating
 PT apoptosis, and for diagnosing and treating cancer
 PS Claim 14; Fig 4; 117pp; English.
 CC The invention relates to Bcl-G polypeptides and nucleic acids encoding
 CC them. The Bcl-G polypeptides can be expressed by standard recombinant
 CC methodology. Bcl-G oligonucleotides (or its anti-sense strand) and Bcl-G
 CC specific antibodies are useful for diagnosing cancer, monitoring cancer
 CC therapy or assessing prognosis of patients with cancer. The Bcl-G
 CC polypeptides are useful for modulating the activity of an oncogenic
 CC polypeptide. They are useful for identifying modulators, for modulating
 CC a level of apoptosis mediated by the Bcl-G polypeptide. A therapeutic
 CC composition comprising the Bcl-G polypeptide, polynucleotide or antibody
 CC is useful for treating a pathology characterized by abnormal cell
 CC proliferation especially cancer. The present sequence represents a
 CC human Bcl-Gs polypeptide.
 CC
 SQ Sequence 252 AA:

Query Match 69.6%; Score 1189.5; DB 22; Length 252;
 Best Local Similarity 93.5%; Pred. No. 6.4e-113;
 Matches 232; Conservative 2; Mismatches 7; Indels 7; Gaps 1;

QY 1 MCSTSGCDLEFIPDDDDNTIEFKILAYTRHHVFKSTPAFPKILRTSLSGRGLGN 60
 1 mcsstsgcdleipdddddntiefkilaaytrhhvfkstpalispkllrtslsgrglgn 60
 DB 61 CSANESWTEVSWPCRNSSQSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKYSAQGRFL 120
 61 csaneswtevwpcrnssqsekaingkksswkaffgvvekedsgstpkysaaggrfl 120
 QY 121 EYQDSHSQOORCLSNVBOCLEHEAVDPKVISTANRYAEIYVSWPPQATQAGFKSKEL 180
 121 eyqdsnsgwscrlsnvboacleheavdpkvistanyaeiyswppqatqagfkskel 180
 DB 121 EYQDSHSQOORCLSNVBOCLEHEAVDPKVISTANRYAEIYVSWPPQATQAGFKSKEL 180
 121 eyqdsnsgwscrlsnvboacleheavdpkvistanyaeiyswppqatqagfkskel 180
 QY 181 FYTEGLSFQLOGHVPYASSSKKDEEQLIAKIVELLKYSQGLERKRLKK-----DKAL 233
 181 fyteglsfqloghvpayassskkdeeqllakivellkysgqlerkdtafipivdtsi 240
 DB 181 fyteglsfqloghvpayassskkdeeqllakivellkysgqlerkdtafipivdtsi 240
 181 fyteglsfqloghvpayassskkdeeqllakivellkysgqlerkdtafipivdtsi 240
 QY 234 MGHFODGL 241
 234 mghfodgl 241
 DB 241 qgfpgdgl 248

RESULT 5
 ID AAB85188 standard; Protein: 328 AA.
 XX AAB85188;
 AC AAB85188;
 DT 07-SEP-2001 (first entry)
 XX Mouse Bcl-G polypeptide.
 DE Mouse Bcl-G polypeptide.
 XX Bcl-G; cancer; cancer therapy; oncogene; apoptosis; cytostatic;
 KW antiapoptotic; chromosome 12p12.3; mouse.
 KM
 XX Mus sp.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 107
 FT /label= unknown
 FT

Query Match	54.48;	Score 929.5;	DB 22;	Length 328
Best Local Similarity	59.58;	Pred. No. 3,66-86.		
Matches 200;	Conservation			

RESULT	6
AA095378	
ID	AA095378 standard; Protein, 151 AA
AC	
XX	
AA095378;	

DT	21-NOV-2001	(first entry)
XX	Human reproductive system related antigen seq ID NO: 4036.	
DE	Human reproductive system related antigen; reproductive system disorder.	
XX		
KW	Human; reproductive system related antigen; reproductive system disorder.	
KW	cancer; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200155320-A2.	
XX		
PD	02-AUG-2001.	
XX		
PF	17-JAN-2001; 2001WO-US01339.	
XX		
PR	31-JAN-2000; 2000US-0179065.	
PR	04-FEB-2000; 2000US-0180658.	
PR	24-FEB-2000; 2000US-0184662.	
PR	02-MAR-2000; 2000US-0186354.	
PR	16-MAR-2000; 2000US-0186370.	
PR	17-MAR-2000; 2000US-0190076.	
PR	18-APR-2000; 2000US-0190076.	
PR	19-MAY-2000; 2000US-0198123.	
PR	07-JUN-2000; 2000US-020515.	
PR	28-JUN-2000; 2000US-0209467.	
PR	30-JUN-2000; 2000US-0214886.	
PR	07-JUL-2000; 2000US-0215135.	
PR	07-JUL-2000; 2000US-0216647.	
PR	11-JUL-2000; 2000US-0216880.	
PR	11-JUL-2000; 2000US-0217487.	
PR	14-JUL-2000; 2000US-0217496.	
PR	26-JUL-2000; 2000US-0218290.	
PR	26-JUL-2000; 2000US-0220963.	
PR	14-AUG-2000; 2000US-0220964.	
PR	14-AUG-2000; 2000US-0224518.	
PR	14-AUG-2000; 2000US-0224519.	
PR	14-AUG-2000; 2000US-0225213.	
PR	14-AUG-2000; 2000US-0225214.	
PR	14-AUG-2000; 2000US-0225266.	
PR	14-AUG-2000; 2000US-0225267.	
PR	14-AUG-2000; 2000US-0225268.	
PR	14-AUG-2000; 2000US-0225447.	
PR	14-AUG-2000; 2000US-0225477.	
PR	14-AUG-2000; 2000US-0225757.	
PR	14-AUG-2000; 2000US-0225758.	
PR	18-AUG-2000; 2000US-0226279.	
PR	22-AUG-2000; 2000US-0226279.	
PR	22-AUG-2000; 2000US-0226681.	
PR	22-AUG-2000; 2000US-0226868.	
PR	23-AUG-2000; 2000US-0227182.	
PR	30-AUG-2000; 2000US-0227009.	
PR	01-SEP-2000; 2000US-0228924.	
PR	01-SEP-2000; 2000US-0229287.	
PR	01-SEP-2000; 2000US-0229343.	
PR	01-SEP-2000; 2000US-0229344.	
PR	05-SEP-2000; 2000US-0229509.	
PR	05-SEP-2000; 2000US-0229513.	
PR	06-SEP-2000; 2000US-0230437.	
PR	06-SEP-2000; 2000US-0230438.	
PR	08-SEP-2000; 2000US-0231242.	
PR	08-SEP-2000; 2000US-0231243.	
PR	08-SEP-2000; 2000US-0231244.	
PR	08-SEP-2000; 2000US-0231413.	
PR	08-SEP-2000; 2000US-0231414.	
PR	08-SEP-2000; 2000US-0232080.	
PR	12-SEP-2000; 2000US-0232081.	
PR	14-SEP-2000; 2000US-0232397.	
PR	14-SEP-2000; 2000US-0232398.	
PR	14-SEP-2000; 2000US-0232399.	
PR	14-SEP-2000; 2000US-0232400.	
PR	14-SEP-2000; 2000US-0232401.	
PR	14-SEP-2000; 2000US-0233063.	

us-09-771-961-2.rag

Thu Jun 20 13:14:55 2002

PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 21-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 25-SEP-2000; 2000US-0235884.
 PR 26-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 27-SEP-2000; 2000US-0236357.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 29-SEP-2000; 2000US-0236802.
 PR 29-SEP-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0238937.
 PR 13-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0246174.
 PR 01-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 08-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249224.
 PR 17-NOV-2000; 2000US-0249225.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 17-NOV-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251890.
 PR 08-DEC-2000; 2000US-0251930.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI: 2001-46570/50.
 DR N-PSDB; AAL01348.
 XX Isolated nucleic acid molecule encoding a reproductive system antigen -
 PT is used in preventing, treating or ameliorating a medical condition.
 PS Claim 11: SEQ ID NO 4036; 1297pp + Sequence Listing: English.
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a protein of the invention.
 CC Sequence 151 AA;
 SO

Query Match 40.5%; Score 691.5; DB 22; Length 151;
 Best Local Similarity 95.0%; Pred. No. 2, 2e-62; Indels 1; Gaps 1;
 Matches 194; Conservative 5; Mismatches 5;

QY 1 MCSTSGCDLEIRPLDDDDINTIEFKILAYTRHHVFKSTPATSPKILRTSRISQRCIGN 69
 Db 10 mcstsgcdleirpldddddntiefkilytrhhvfkscpalisfkllrtsrtsqrgl 129
 QY 61 CSANSTREVSMPCRNSQSEKAINIGKSSMKAFEGVEKEQSOPAKVSNOGORTL 129
 Db 70 csanstevsmpcrnsqsekaingkssmkafegvexedsqspakvsagqrcl 129
 QY 121 EYODSHSQMSRCLSNVEQCL 141
 Db 130 eyodshss-gqylsnveqvl 149

RESULT 7
 ABB27961 standard; Peptide; 129 AA.
 ID ABB27961
 AC ABB27961;
 DT 01-FEB-2002 (first entry)
 DE Human peptide #612 encoded by breast cell single exon nucleic acid probe.
 DE Human; microarray: single exon probe; gene expression; breast;
 KW disease; cancer.
 OS Homo sapiens.
 XX WO200157271-R2.
 PN 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-US00662.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632267.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 PR (MOLE-) MOLECULAR DYNAMICS INC.

...nucleic acid probes for ...
...relates to a spatially ...
...English.

PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0633265.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0234687.
 PR 04-OCT-2000; 2000US-0236359.
 XX 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483447/52.
 DR
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 XX Claim 27; SEQ ID NO 25768; 63pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human
 CC liver. The single exon nucleic acid probes may be used for predicting
 CC fetal liver. The displaying gene expression in samples derived from human
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp://wipo.int/pub/published_pat_sequences
 XX
 XX Sequence 129 at:

16 DDDLNTIEFKILAYTRHVFKSTPALESPKLLRTRSTSOBCT 0,

Query Match	40.1%	Score 686,	DB 22:	Length 129,
Best Local Similarity	100.0%	Pred. No. 6, 3-e-62,		
Matches 129;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	16	DDDLNTIEKILAYTRHHVFKSTPALFSEKILRSTSQRLGNCNSANESWTFVSPDC		
Db	1	dddnltlefkilaytrthvfkstpalrfskllrtstsqrlgncansawtfvswpcr		
Qy	76	NSQSEKAININGKKSQKMAFGVVEKEESQSTPAKYSAGQORTLEQDHSQMSQSRCLS		
Db	61	nsqsekeainlqkkkskwaifgvvekedsqstprakvsagqgrtleqdsqsgsqsrcls		
Qy	136	NVBDQLEHE 144		
Db	121	nveqgclehe 120		

RESULT	9
ABBI8598	
ID	ABBI8598 standard; Protein; 129 AA
XX	
AC	ABBI8598.

2002 (first entry)

cardiac gene expression: heart gene expression

congenital heart disease; hypertension; cardiovascular disease; vascular system

homo sapiens.

WC200157274-A2

09-AUG-2001.

30-JAN-2007: 2001ms

04-FEB-2000: 2000--

30-JUN-2000; 2000US-0207456

2000; 2000US-0632366

Thu Jun 20 13:14:55 2002

us-09-771-961-2.1rag

Page 7

21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI: 2001-488899/53.
Single exon nucleic acid probes for analyzing gene expression in human brains -
Claim 15; SEQ ID NO 20368; 530pp; English.
The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart (see ABA21535-ABA41305). The present sequence is a protein encoded by one such probe. The probes may be used for predicting, measuring and displaying the gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosticating diseases of human heart and vascular system e.g. cardiovascular heart disease, hypertension, cardiac arrhythmias and did not form part of the printed specification, but was obtained in electronic format directly from WPI at tcp.wpi.int/pub/published_pct_sequences.
Sequence 129 AA:
40.1%; Score 686; DB 22; Length 129; 0;
Query Match Similarity 100.0%; Pred. No. 6.3e-62; Indels 0; Gaps 0;
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;
Matches 129; Conservative
16 DDDLTIEFKILAYTRHHVFKSTPALFSPKILRTSLISORGLGNCANSEWTEVSWPCR 75
1 dddlnliefkllaytrhhvfkstpalfspkllrtslisqrglncanesevtevwpcr 60
76 NSQSSSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSNOGORTLEYODSHSOQMSRCLS 135
1 nsgsssekainlgkkskwaffgvvekedsgstpakvsnoqgortleyodshsqmsrccls 120
76 NSQSSSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSNOGORTLEYODSHSOQMSRCLS 135
1 nsgsssekainlgkkskwaffgvvekedsgstpakvsnoqgortleyodshsqmsrccls 120
136 NVEQCLEHE 144
1 nveqclehe 129
121 nveqclehe 129
RESULT 10
ID AAM53929 standard; Protein: 129 AA.
AAM53929:
05-NOV-2001 (first entry)
Human brain expressed single exon probe encoded protein SEQ ID NO: 26034.
Human brain expressed exon; gene expression analysis: probe;
microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
epilepsy; cancer.
Homo sapiens.
WO200157275-A2.
09-AUG-2001.
30-JAN-2001; 2001WO-US00667.
04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632365.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI: 2001-483446/52.
Single exon nucleic acid probes for analyzing gene expression in human brains -
Example 4; SEQ ID NO: 26034; 650pp + Sequence listing; English.
The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is a protein encoded by one of the probes of the invention.
Sequence 129 AA:
40.1%; Score 686; DB 22; Length 129; 0;
Query Match Similarity 100.0%; Pred. No. 6.3e-62; Indels 0; Gaps 0;
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;
Matches 129; Conservative
16 DDDLTIEFKILAYTRHHVFKSTPALFSPKILRTSLISORGLGNCANSEWTEVSWPCR 75
1 dddlnliefkllaytrhhvfkstpalfspkllrtslisqrglncanesevtevwpcr 60
76 NSQSSSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSNOGORTLEYODSHSOQMSRCLS 135
1 nsgsssekainlgkkskwaffgvvekedsgstpakvsnoqgortleyodshsqmsrccls 120
136 NVEQCLEHE 144
1 nveqclehe 129
121 nveqclehe 129
RESULT 11
ID AAM6317 standard; Protein: 129 AA.
AAM6317:
06-NOV-2001 (first entry)
Human bone marrow expressed probe encoded protein SEQ ID NO: 26623.
Human bone marrow expressed exon; gene expression analysis: probe;
microarray; cancer; leukaemia; lymphoma; myeloma.
Homo sapiens.
WO200157276-A2.
09-AUG-2001.
30-JAN-2001; 2001WO-US00668.
04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632365.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.

	Conservative	Mismatches	Indels	Gaps
Qy	16	0	0	0
Db	1	0	0	0
Oy	76	0	0	0
Db	61	0	0	0
Oy	136	0	0	0
Db	121	0	0	0

RESULT	12
AA14186	
ID	AA14186
XX	standard; Protein; 129 AA
AC	
XX	AA14186;

AA Peptide #620 (first entry)
DE
XX
XX Probe; human; microarray; gene expression
XX cervical cancer.
XX Homo sapiens.
XX cervical epithelial cell;

XX		humo sapiens.
PN		MO2001572Z-A2.
PX		
PD		09-AUG-2001.
XX		
PF		
PR		30-JAN-2001; 2001MO-US00670.
PR		
PR	04-FEB-2000;	2000US-0180312.
PR	26-MAY-2000;	2000US-0207456.
PR	30-JUN-2000;	2000US-0207456.
PR	03-AUG-2000;	2000US-0608408.
PR	21-SEP-2000;	2000US-0632366.
PR	27-SEP-2000;	2000US-0234687.
XX	04-OCT-2000;	2000US-0236359.
XX		2000GB-0024263.
PA	(MOLE-)	MOLECULAR DYNAMICS INC.
XX		
PI		
DR	Penn SG, Hanzel DK, Chen W,	Rank DR.
XX	WPI; 2001-488901/53.	

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
XX
PS
XX
XX
CC
CC
CC (SENp: see A4110068, A4128459). The present sequence is a peptide encoded
CC by one such probe. The SENps are derived from human HeLa cells. The SEN
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probe can be used for
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO
CC at [ftp.wipo.int/pub/published_pct-sequences](http://wipo.int/pub/published_pct-sequences).
XX
XX
S0 Sequence 129 Aa;

Query Match	40.1%;	Score 686;	DB 22;	Length 129;
Best Local Similarity	100.0%;	Pred. No. 6.3e-62;		
Matches 129;	Conservative	0;	Mismatches	
16 Percent				

	0:	Gaps
Db	1 dddltltfiskllaylttnhkfcpalrfspklltrtslsqgylncasnewtlevsmpdr	75
Oy	76 NQSEKAININGKKSSKKKAFEGVEEKEDSDSPAKYASAGORTLEYQDSHSQOQMSRLS	60
Db	61 nsqsekaalilgkkskswalfgvekedsgstcpakysagqgqrltleygdshsqwtswrlcis	120
Oy	136 NVEQCLEHE 144	
Db	121 nveqclehe 129	

RESULT	13
AAM26596	
ID	AAM26596
XX	standard; Protein; 129 AA
AC	
AY	AAM26596;

XX 17-OCT-2001 (first entry)
DE
XX
XX Peptide #633 encoded by probe for measuring placental gene expression
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder.
XX
XX Homo sapiens.
XX

PD	09-AUG-2001.
XX	
XX	30-JAN-2001; 2001KO-US00663
XX	
PR	04-FEB-2000; 2000US-0180312
PR	26-MAR-2000; 2000US-0180312
PR	30-JUN-2000; 2000US-0207456
PR	03-AUG-2000; 2000US-0608408
PR	21-SEP-2000; 2000US-0633366
PR	21-SEP-2000; 2000US-0234687
PR	04-OCT-2000; 2000US-0236587
PA	(MOLE - MORT); 2000GB-0024263.

	PI	XX	XX	DR	XX	PT
	Penn SG,	Hanzel DK,	Chen W,	Rank DR,	WPI, 2001-488897/53.	Human genome-3

... exon nucleic acid probes useful for

RESULT 14
AAM01922
1000 standard; Protein; 129 AA

XX
09-OCT-2001 (first entry)

Abb. Probe; human; breast disease, non-inflammatory disease; proliferative breast disease, inflammatory disease; breast cancer

W02001572/0-

XX
RE 29-JAN-2001; 2001WO-US006661.

PR 20 JUN - 2000; 2000US - 0608408
PR 30 - JUN - 2000; 2000US - 0632366

27-SEP-2000; 2000GB-0024263
04-OCT-2000; PR

XX	SG	XX	SG
Hanzel DK,	Chen W,	Rank DK	

XX 3' single exon nucleic acid probe used to measure

Claim 27; SEQ ID No 10002, 10003, 10004, 10005, 10006, 10007, 10008, 10009, 10010, 10011, 10012, 10013, 10014, 10015, 10016, 10017, 10018, 10019, 10020, 10021, 10022, 10023, 10024, 10025, 10026, 10027, 10028, 10029, 10030, 10031, 10032, 10033, 10034, 10035, 10036, 10037, 10038, 10039, 10040, 10041, 10042, 10043, 10044, 10045, 10046, 10047, 10048, 10049, 10050, 10051, 10052, 10053, 10054, 10055, 10056, 10057, 10058, 10059, 10060, 10061, 10062, 10063, 10064, 10065, 10066, 10067, 10068, 10069, 10070, 10071, 10072, 10073, 10074, 10075, 10076, 10077, 10078, 10079, 10080, 10081, 10082, 10083, 10084, 10085, 10086, 10087, 10088, 10089, 10090, 10091, 10092, 10093, 10094, 10095, 10096, 10097, 10098, 10099, 10100, 10101, 10102, 10103, 10104, 10105, 10106, 10107, 10108, 10109, 10110, 10111, 10112, 10113, 10114, 10115, 10116, 10117, 10118, 10119, 10120, 10121, 10122, 10123, 10124, 10125, 10126, 10127, 10128, 10129, 10130, 10131, 10132, 10133, 10134, 10135, 10136, 10137, 10138, 10139, 10140, 10141, 10142, 10143, 10144, 10145, 10146, 10147, 10148, 10149, 10150, 10151, 10152, 10153, 10154, 10155, 10156, 10157, 10158, 10159, 10160, 10161, 10162, 10163, 10164, 10165, 10166, 10167, 10168, 10169, 10170, 10171, 10172, 10173, 10174, 10175, 10176, 10177, 10178, 10179, 10180, 10181, 10182, 10183, 10184, 10185, 10186, 10187, 10188, 10189, 10190, 10191, 10192, 10193, 10194, 10195, 10196, 10197, 10198, 10199, 10200, 10201, 10202, 10203, 10204, 10205, 10206, 10207, 10208, 10209, 10210, 10211, 10212, 10213, 10214, 10215, 10216, 10217, 10218, 10219, 10220, 10221, 10222, 10223, 10224, 10225, 10226, 10227, 10228, 10229, 10230, 10231, 10232, 10233, 10234, 10235, 10236, 10237, 10238, 10239, 10240, 10241, 10242, 10243, 10244, 10245, 10246, 10247, 10248, 10249, 10250, 10251, 10252, 10253, 10254, 10255, 10256, 10257, 10258, 10259, 10260, 10261, 10262, 10263, 10264, 10265, 10266, 10267, 10268, 10269, 10270, 10271, 10272, 10273, 10274, 10275, 10276, 10277, 10278, 10279, 10280, 10281, 10282, 10283, 10284, 10285, 10286, 10287, 10288, 10289, 10290, 10291, 10292, 10293, 10294, 10295, 10296, 10297, 10298, 10299, 10300, 10301, 10302, 10303, 10304, 10305, 10306, 10307, 10308, 10309, 10310, 10311, 10312, 10313, 10314, 10315, 10316, 10317, 10318, 10319, 10320, 10321, 10322, 10323, 10324, 10325, 10326, 10327, 10328, 10329, 10330, 10331, 10332, 10333, 10334, 10335, 10336, 10337, 10338, 10339, 10340, 10341, 10342, 10343, 10344, 10345, 10346, 10347, 10348, 10349, 10350, 10351, 10352, 10353, 10354, 10355, 10356, 10357, 10358, 10359, 10360, 10361, 10362, 10363, 10364, 10365, 10366, 10367, 10368, 10369, 10370, 10371, 10372, 10373, 10374, 10375, 10376, 10377, 10378, 10379, 10380, 10381, 10382, 10383, 10384, 10385, 10386, 10387, 10388, 10389, 10390, 10391, 10392, 10393, 10394, 10395, 10396, 10397, 10398, 10399, 10400, 10401, 10402, 10403, 10404, 10405, 10406, 10407, 10408, 10409, 10410, 10411, 10412, 10413, 10414, 10415, 10416, 10417, 10418, 10419, 10420, 10421, 10422, 10423, 10424, 10425, 10426, 10427, 10428, 10429, 10430, 10431, 10432, 10433, 10434, 10435, 10436, 10437, 10438, 10439, 10440, 10441, 10442, 10443, 10444, 10445, 10446, 10447, 10448, 10449, 10450, 10451, 10452, 10453, 10454, 10455, 10456, 10457, 10458, 10459, 10460, 10461, 10462, 10463, 10464, 10465, 10466, 10467, 10468, 10469, 10470, 10471, 10472, 10473, 10474, 10475, 10476, 10477, 10478, 10479, 10480, 10481, 10482, 10483, 10484, 10485, 10486, 10487, 10488, 10489, 10490, 10491, 10492, 10493, 10494, 10495, 10496, 10497, 10498, 10499, 10500, 10501, 10502, 10503, 10504, 10505, 10506, 10507, 10508, 10509, 10510, 10511, 10512, 10513, 10514, 10515, 10516, 10517, 10518, 10519, 10520, 10521, 10522, 10523, 10524, 10525, 10526, 10527, 10528, 10529, 10530, 10531, 10532, 10533, 10534, 10535, 10536, 10537, 10538, 10539, 10540, 10541, 10542, 10543, 10544, 10545, 10546, 10547, 10548, 10549, 10550, 10551, 10552, 10553, 10554, 10555, 10556, 10557, 10558, 10559, 10560, 10561, 10562, 10563, 10564, 10565, 10566, 10567, 10568, 10569, 10570, 10571, 10572, 10573, 10574, 10575, 10576, 10577, 10578, 10579, 10580, 10581, 10582, 10583, 10584, 1058

CC The present inventory is a copy of the original. The present sequence is 1-10.
CC (see AAI00010-AAI10067).
CC

```
Query Match      40.1%; Score 686; DB 22; Length 129;          0
Best Local Similarity 100.0%; Pred. No. 6.3e-62; Indels 0; Gaps 0
Matches 129; Conservative 0; Mismatches 0; SNPCF 75
```

1 aaalncicn-----

db 61 nsqsssekallllyknc...

Db 121 nveqclehe 129

RESULT 10
ABG26266

AC	XX	0000 (first entry)

chromosome mapping; gene mapping; genetic disorder.
XX
XX

aa
05 Homo sapiens.

11-001 PD XX

AA 2000US-0340227
PR 2000US-0649167
31-MAR-2000;
2000

XX Liu C, Tang Y
XX PT, Tang Y

DR N-PSDB; AAS90453.
DR

PT responsible for genetic disorders -

Claim 20, comprising the limitation "isolated polynucleotide"

CC polymerase chain reaction (PCR) primers, oligomers, and sequences. (2)

CC polypeptide (II)

CC

CC

1

CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful for medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations in
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. Abc00010-Abc30377 represent novel human
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 178 AA;

RY Match 35.6%; Score 608; DB 22; Length 178;
 Best Local Similarity 100.0%; Pred. No. 9.6e-54;
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 211 KIVELKYSQDLEKTKKALMGHFDGLSYVFKTTDOYLMGVDPKGESEVKAQGF 270
 Db 2 KIVELKYSQDLEKTKKALMGHFDGLSYVFKTTDOYLMGVDPKGESEVKAQGF 270
 QY 271 KALVYDVYAKLAIIDNHPNRYLGFGRKYLKENSPIIDQHGMEKILGISHEVD 327
 Db 62 kaalvldvtakilaalidnphmrvlygfgtkylkenfswlqhgweklqishvevd 118

Search completed: June 19, 2002, 16:34:38
 Job time: 131 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 16:34:38 ; Search time 53.19 Seconds
(without alignments)
526.238 Million cell updates/sec

Title: US-09-771-961-4
Perfect score: 1316
Sequence: 1 MCRSTGCDLEIPDDDLN.....IPVDTSTGTFPGQIMACI 252

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: A-Geneset_032802:*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
23: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1316	100.0	252	AAE07041	Human BCL-X-like p
2	1316	100.0	252	AAE07041	Human BCL-X-like p
3	1189.5	90.4	327	AAE07040	Human BCL-X-like p
4	1189.5	90.4	327	AAE07040	Human BCL-X-like p
5	691.5	52.5	151	AAE05378	Human BCL-X-like p
6	686	52.1	129	AAE05378	Human BCL-X-like p
7	686	52.1	129	AAE05378	Human BCL-X-like p
8	686	52.1	129	AAE05378	Human BCL-X-like p
9	686	52.1	129	AAE05378	Human BCL-X-like p
10	686	52.1	129	AAE05378	Human BCL-X-like p
11	686	52.1	129	AAE05378	Human BCL-X-like p

12	686	52.1	129	AAE05378	Human BCL-X-like p
13	686	52.1	129	AAE05378	Human BCL-X-like p
14	501	38.1	328	AAE05378	Human BCL-X-like p
15	93	7.1	1250	AAE05378	Human BCL-X-like p
16	91.5	7.0	772	AAE05378	Human BCL-X-like p
17	91.5	7.0	1075	AAE05378	Human BCL-X-like p
18	91.5	7.0	1075	AAE05378	Human BCL-X-like p
19	90	6.8	660	AAE05378	Human BCL-X-like p
20	89.5	6.8	696	AAE05378	Human BCL-X-like p
21	89	6.8	932	AAE05378	Human BCL-X-like p
22	89	6.8	1252	AAE05378	Human BCL-X-like p
23	88.5	6.7	178	AAE05378	Human BCL-X-like p
24	88.5	6.7	1369	AAE05378	Human BCL-X-like p
25	88.5	6.7	1369	AAE05378	Human BCL-X-like p
26	86	6.5	639	AAE05378	Human BCL-X-like p
27	85.5	6.5	1311	AAE05378	Human BCL-X-like p
28	85	6.5	635	AAE05378	Human BCL-X-like p
29	85	6.5	808	AAE05378	Human BCL-X-like p
30	83	6.3	277	AAE05378	Human BCL-X-like p
31	83	6.3	331	AAE05378	Human BCL-X-like p
32	83	6.3	634	AAE05378	Human BCL-X-like p
33	83	6.3	1291	AAE05378	Human BCL-X-like p
34	82.5	6.3	1148	AAE05378	Human BCL-X-like p
35	82.5	6.3	1148	AAE05378	Human BCL-X-like p
36	81.5	6.2	250	AAE05378	Human BCL-X-like p
37	81.5	6.2	250	AAE05378	Human BCL-X-like p
38	81.5	6.2	253	AAE05378	Human BCL-X-like p
39	81.5	6.2	253	AAE05378	Human BCL-X-like p
40	81.5	6.2	1094	AAE05378	Human BCL-X-like p
41	81.5	6.2	807	AAE05378	Human BCL-X-like p
42	81	6.2	1029	AAE05378	Human BCL-X-like p
43	80.5	6.1	777	AAE05378	Human BCL-X-like p
44	80.5	6.1	932	AAE05378	Human BCL-X-like p

ALIGNMENTS

RESULT 1	AAE07041	standard; Protein: 252 AA.
XX	AAE07041	
AC	AAE07041	
XX	23-OCT-2001 (first entry)	
DT	XX	Human BCL-X-like protein #2.
XX	XX	Human: BCL-X-like protein; physiological disorder.
DE	XX	
XX	XX	
KW	XX	Homo sapiens.
OS	XX	
XX	XX	WO200157213-A2.
FN	XX	
XX	XX	09-AUG-2001.
PD	XX	
XX	XX	31-JAN-2001; 2001WO-US03446.
PF	XX	
XX	XX	04-FEB-2000; 2000US-0180412.
PR	XX	
XX	XX	(LEXI-) LEXICON GENETICS INC.
PA	XX	
PI	XX	Donoho G, Hilbun E, Turner CA, Friedrich G, Abuin A, Zambrowicz B;
PI	XX	Sands AT;
XX	XX	WPI: 2001-488882/53.
DR	XX	N-PSDB: AAD13336.
XX	XX	Novel isolated human BCL-X-like polynucleotide, useful in therapeutic,
PT	XX	diagnostic and pharmacogenic applications
XX	XX	Claim 5; page 32; 33pp; English.

Peptide #633 encod
Peptide #604 encod
Mouse Bcl-G polyp
Amino acid sequenc
Homo sapiens DH130
Human secreted pro
Human DL162.2 pro
Human polypeptide
Drosophila melanog
Mouse Nope (neighb
Mouse Nope (neighb
Novel human diagno
Human secreted pro
Human secreted pro
Human polypeptide
Drosophila melanog
Human polypeptide
Human cytoplasmic
Breast and ovarian
Human polypeptide
Human polypeptide
Human AC15 protein
Human AC15 protein
Human OREX ORF2721
Human protein SEQ
Human protein sequ
Human CGI-97 prote
Human protein SEQ
Human homer intera
Human prp04. Homo
Amino acid sequenc
Homo sapiens BAR1
Drosophila melanog

XX The present sequence is human BCL-X-like protein.
CC The BCL-X-like polynucleotides are useful in therapeutic, diagnostic
CC and pharmacogenetic applications. They are useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic manifestations
CC perturbing the normal function of protein in the body and also for
CC treating physiological disorders and diseases. The BCL-X-like
CC polynucleotides are useful in conjunction with polymerase chain
CC reaction to screen libraries, isolate clones, to prepare cloning
CC and sequencing templates and as hybridisation probes for assessing
CC gene expression patterns.
XX
SQ Sequence 252 AA:

Query Match 100.0%; Score 1316; DB 22; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.7e-131;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCSTSGCDLEIPDDDDLTIEFKILAYTRHVFSTPALFSPKLLRTSLSGRIGN 60
F 1 mctsgcdleipdddddltiefkilaaytrhvfstpalfspkllrtslsgrlgn 60
QY 61 CSANESWTEVSWPCRNSSSEKAINLGKSSWKAFEGVEKEDSOSTPAKVASAGQRTL 120
Db 61 csanewtevwpcrnsssekainlgksswkafegvekedsgstpakvasagqrtl 120
QY 121 EYODSHSOQSRCLSNVEQCLEHEAVDPKVISIANRYAEIYSWPPQATQAGFSKET 180
Db 121 eyodshsqsrclsnveqcleheavdpkvisianryaeiyswppqatqagfsket 180
QY 181 FVTEGLSFQLOGHVPVASSSKKDEEOILAKIVELLYSGDLEKRTAFIPPLVDTSI 240
Db 181 fvtelgsfqlqghvpassskkdeeqilakivellykysgdlerktafipplvdtsi 240
QY 241 QGFPODGLMACI 252
Db 241 qgfpgdglmaci 252

RESULT 2

AAB85167
ID AAB85167 standard; Protein; 252 AA.

XX AAB85167;

DT 07-SEP-2001 (first entry)

XX Human BCL-Gs polypeptide.

XX BCL-G; cancer; cancer therapy; oncogene; apoptosis; BCL-Gs; cytostatic;
XX antiapoptotic; chromosome 12p12.3; human.

XX Homo sapiens.

XX WO200144282-A2.

XX 21-JUN-2001.

XX 13-DEC-2000; 2000MO-US33793.

XX 14-DEC-1999; 99US-0461641.

XX (BURN-) BURNHAM INST.

XX Reed JC, Godzik A;

XX WPI; 2001-398125/42.

XX N-PSDB; AAH22583.

XX Novel polynucleotide encoding BCL-G polypeptide, useful for modulating
XX apoptosis, and for diagnosing and treating cancer

PS Claim 14; Fig 4; 11Pp; English.

XX The invention relates to BCL-G polypeptides and nucleic acids encoding
CC them. The BCL-G polypeptides can be expressed by standard recombinant
CC methodology. BCL-G oligonucleotides (or its anti-sense strand) and BCL-G
CC specific antibodies are useful for diagnosing cancer, monitoring cancer
CC therapy or assessing prognosis of patients with cancer. The BCL-G
CC polypeptides are useful for modulating the activity of an oncogenic
CC a level of apoptosis mediated by the BCL-G polypeptide, for modulating
CC composition comprising the BCL-G polypeptide, polynucleotide or antibody
CC proliferation especially cancer. The present sequence represents a
CC human BCL-Gs polypeptide.
XX
SQ Sequence 252 AA:

Query Match 100.0%; Score 1316; DB 22; Length 252;
Best Local Similarity 100.0%; Pred. No. 9.7e-131;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCSTSGCDLEIPDDDDLTIEFKILAYTRHVFSTPALFSPKLLRTSLSGRIGN 60
F 1 mctsgcdleipdddddltiefkilaaytrhvfstpalfspkllrtslsgrlgn 60
QY 61 CSANESWTEVSWPCRNSSSEKAINLGKSSWKAFEGVEKEDSOSTPAKVASAGQRTL 120
Db 61 csanewtevwpcrnsssekainlgksswkafegvekedsgstpakvasagqrtl 120
QY 121 EYODSHSOQSRCLSNVEQCLEHEAVDPKVISIANRYAEIYSWPPQATQAGFSKET 180
Db 121 eyodshsqsrclsnveqcleheavdpkvisianryaeiyswppqatqagfsket 180
QY 181 FVTEGLSFQLOGHVPVASSSKKDEEOILAKIVELLYSGDLEKRTAFIPPLVDTSI 240
Db 181 fvtelgsfqlqghvpassskkdeeqilakivellykysgdlerktafipplvdtsi 240
QY 241 QGFPODGLMACI 252
Db 241 qgfpgdglmaci 252

RESULT 3

AAE07040
ID AAE07040 standard; Protein; 327 AA.

XX AAE07040;

DT 23-OCT-2001 (first entry)

XX Human BCL-X-like protein #1.

XX Human; BCL-X-like protein; therapy; physiological disorder.

XX Homo sapiens.

XX WO200157213-A2.

XX 09-AUG-2001.

XX 31-JAN-2001; 2001MO-US03446.

XX 04-FEB-2000; 2000US-0180412.

XX (LEXI-) LEXICON GENETICS INC.

XX Donoho G, Hilbun E, Turner CA, Friedrich G, Abuin A, Zambrowicz B;
XX Sands AT;

XX WPI; 2001-488882/53.
XX N-PSDB; AAD13235.

PT Novel isolated human BCL-X-like polynucleotide, useful in therapeutic,
 PR diagnostic and pharmacogenic applications
 XX
 PS Claim 4; Page 30-31; 33pp; English.

XX The present sequence is human BCL-X-like protein.
 CC The BCL-X-like polynucleotides are useful in therapeutic, diagnostic
 CC and pharmacogenic applications. They are useful for screening drugs
 CC effective in the treatment of symptomatic or phenotypic manifestations
 CC perturbing the normal function of protein in the body and also for
 CC treating physiological disorders and diseases. The BCL-X-like
 CC polynucleotides are useful in conjunction with polymerase chain
 CC reaction to screen libraries, isolate clones, to prepare cloning
 CC and sequencing templates and as hybridisation probes for assessing
 CC gene expression patterns.

CC Sequence 327 AA;

SO Query Match 90.4%; Score 1189.5; DB 22; Length 327;

Best Local Similarity 93.5%; Pred. No. 3.5e-117; Indels 7; Gaps 1;
 Matches 232; Conservative 2; Mismatches 7;

QY 1 MCSTSGCDLEIPDDDDLNTIEFKILAYTRHHVFKSTPALFSPKILTRSLSGRLGN 60
 1 mctsgcdleipdddddntiefkllaytrhhvfksstpalfspklltrslsgrlgn 60
 DB 61 CSANESMTSEVSPKNSQSEKAINLGKRSKSWKAFEGVEKEDSOSTPAKVSAGQRTL 120
 61 csanestevspkncsqsekaingkrskswkafegvekedsgstpakvsagqrtl 120
 DB 121 EYDSSHQMSRCLSNVEQCLEHEAVDPKVISIANRVAEIVSWPPQATQAGFKSKEI 180
 121 eydshsqmsrclsnveqcleheavdpkvisianrvaelyswppqatqagfkskel 180
 DB 121 eydshsqmsrclsnveqcleheavdpkvisianrvaelyswppqatqagfkskel 180
 QY 181 FVTEGSLFQLOGHPVAVASSSKKDEEOILAKIVELKYSQDLEKRDATFIPPLVDTSI 240
 181 fvtelgslfqloghpvavassskkdeeoilakivellkysgdlerklk-----dkel 233
 DB 181 fvtelgslfqloghpvavassskkdeeoilakivellkysgdlerklk-----dkel 233
 QY 241 QGFPODGL 248
 241 qgfpo dgl 241
 DB 234 mghfdqgl 241

RESULT 4

AAB85166 standard; Protein: 327 AA.

AAB85166;

XX 07-SEP-2001 (first entry)
 DT Human Bcl-G1 polypeptide.
 XX
 DE Human Bcl-G1 polypeptide.
 KW Bcl-G1; cancer; cancer therapy; oncogene; apoptosis; Bcl-G1; cytostatic;
 KW antiproliferative; chromosome 12p12.3; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200144282-A2.
 PD 21-JUN-2001.
 XX
 PF 13-DEC-2000; 2000MO-US33793.
 XX
 PR 14-DEC-1999; 99US-0461641.
 XX
 PA (BURN-) BURHAM INST.
 XX
 PI Reed JC, Godzik A;
 XX
 XX WPT: 2001-398125/42.
 DR N-PSDB; AAH22582.

XX Novel polynucleotide encoding Bcl-G polypeptide, useful for modulating
 PR apoptosis, and for diagnosing and treating cancer
 XX
 PS Claim 14; Fig 2; 117pp; English.

XX The invention relates to Bcl-G polypeptides and nucleic acids encoding
 CC them. The Bcl-G polypeptides can be expressed by standard recombinant
 CC methodology. Bcl-G oligonucleotides (or its anti-sense strand) and Bcl-G
 CC specific antibodies are useful for diagnosing cancer, monitoring cancer
 CC therapy or assessing prognosis of patients with cancer. The Bcl-G
 CC polypeptides are useful for modulating the activity of an oncogenic
 CC polypeptide. They are useful for identifying modulators for modulating
 CC a level of apoptosis mediated by the Bcl-G polypeptide. A therapeutic
 CC composition comprising the Bcl-G polypeptide, polynucleotide or antibody
 CC is useful for treating a pathology characterized by abnormal cell
 CC proliferation especially cancer. The present sequence represents a
 CC human Bcl-G polypeptide.

CC Sequence 327 AA;

SO Query Match 90.4%; Score 1189.5; DB 22; Length 327;

Best Local Similarity 93.5%; Pred. No. 3.5e-117; Indels 7; Gaps 1;
 Matches 232; Conservative 2; Mismatches 7;

QY 1 MCSTSGCDLEIPDDDDLNTIEFKILAYTRHHVFKSTPALFSPKILTRSLSGRLGN 60
 1 mctsgcdleipdddddntiefkllaytrhhvfksstpalfspklltrslsgrlgn 60
 DB 61 CSANESMTSEVSPKNSQSEKAINLGKRSKSWKAFEGVEKEDSOSTPAKVSAGQRTL 120
 61 csanestevspkncsqsekaingkrskswkafegvekedsgstpakvsagqrtl 120
 DB 121 EYDSSHQMSRCLSNVEQCLEHEAVDPKVISIANRVAEIVSWPPQATQAGFKSKEI 180
 121 eydshsqmsrclsnveqcleheavdpkvisianrvaelyswppqatqagfkskel 180
 DB 121 eydshsqmsrclsnveqcleheavdpkvisianrvaelyswppqatqagfkskel 180
 QY 181 FVTEGSLFQLOGHPVAVASSSKKDEEOILAKIVELKYSQDLEKRDATFIPPLVDTSI 240
 181 fvtelgslfqloghpvavassskkdeeoilakivellkysgdlerklk-----dkel 233
 DB 181 fvtelgslfqloghpvavassskkdeeoilakivellkysgdlerklk-----dkel 233
 QY 241 QGFPODGL 248
 241 qgfpo dgl 241
 DB 234 mghfdqgl 241

RESULT 5

AAM95378 standard; Protein: 151 AA.

AAM95378;

XX 21-NOV-2001 (first entry)
 DT Human reproductive system related antigen SEQ ID NO: 4036.
 XX
 DE Human reproductive system related antigen; reproductive system disorder;
 KW Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200155320-A2.
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01339.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 XX
 PR 04-FEB-2000; 2000US-0180628.
 XX
 PR 24-FEB-2000; 2000US-0184664.
 XX
 PR 02-MAR-2000; 2000US-0186350.
 XX
 PR 16-MAR-2000; 2000US-0189874.

PR	13-OCT-2000	2000US-0239937	
PR	20-OCT-2000	2000US-0240960	
PR	20-OCT-2000	2000US-0241221	
PR	20-OCT-2000	2000US-0241785	
PR	20-OCT-2000	2000US-0241786	
PR	20-OCT-2000	2000US-0241787	
PR	20-OCT-2000	2000US-0241808	
PR	20-OCT-2000	2000US-0241809	
PR	01-NOV-2000	2000US-0241826	
PR	08-NOV-2000	2000US-0244617	
PR	08-NOV-2000	2000US-0246474	
PR	08-NOV-2000	2000US-0246475	
PR	08-NOV-2000	2000US-0246476	
PR	08-NOV-2000	2000US-0246477	
PR	08-NOV-2000	2000US-0246523	
PR	08-NOV-2000	2000US-0246524	
PR	08-NOV-2000	2000US-0246525	
PR	08-NOV-2000	2000US-0246526	
PR	08-NOV-2000	2000US-0246527	
PR	08-NOV-2000	2000US-0246528	
PR	08-NOV-2000	2000US-0246532	
PR	08-NOV-2000	2000US-0246609	
PR	08-NOV-2000	2000US-0246610	
PR	08-NOV-2000	2000US-0246611	
PR	17-NOV-2000	2000US-0246613	
PR	17-NOV-2000	2000US-0249207	
PR	17-NOV-2000	2000US-0249208	
PR	17-NOV-2000	2000US-0249209	
PR	17-NOV-2000	2000US-0249210	
PR	17-NOV-2000	2000US-0249211	
PR	17-NOV-2000	2000US-0249212	
PR	17-NOV-2000	2000US-0249213	
PR	17-NOV-2000	2000US-0249214	
PR	17-NOV-2000	2000US-0249215	
PR	17-NOV-2000	2000US-0249216	
PR	17-NOV-2000	2000US-0249217	
PR	17-NOV-2000	2000US-0249218	
PR	17-NOV-2000	2000US-0249244	
PR	17-NOV-2000	2000US-0249264	
PR	17-NOV-2000	2000US-0249265	
PR	17-NOV-2000	2000US-0249266	
PR	17-NOV-2000	2000US-0249297	
PR	17-NOV-2000	2000US-0249300	
PR	01-DEC-2000	2000US-0250160	
PR	01-DEC-2000	2000US-0250391	
PR	05-DEC-2000	2000US-0251030	
PR	05-DEC-2000	2000US-0251988	
PR	06-DEC-2000	2000US-0256577	
PR	08-DEC-2000	2000US-0251479	
PR	08-DEC-2000	2000US-0251856	
PR	08-DEC-2000	2000US-0251868	
PR	08-DEC-2000	2000US-0251869	
PR	08-DEC-2000	2000US-0251989	
PR	11-DEC-2000	2000US-0254097	
PR	05-JAN-2001	2001US-0259678	
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Rosen CA, Barash SC, Ruben SM,		
XX	WPI: 2001-465570/50.		
XX	N-PSDB; AAL01346.		
DR			
XX			
PT	Isolated nucleic acid molecule e		
PR	is used in preventing, treating		
XX	Claim 11: SEQ ID NO 4036; 1297pp		
XX			
CC	The present invention provides t		
CC	number of human reproductiv		

The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used

in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a protein of the invention.

Sequence 151 AA:

Query Match 52.5%; Score 691.5; DB 22; Length 151;
Best Local Similarity 95.0%; Pred. No. 5.6e-65;
Matches 134; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

1 MCGSTGQDLEIPDDDDNTIEFKILAYTRHHVFKSPALFSPKILRTSLSGRLGN 60
10 mcsstsgdleeipdddddntiefkilytrhhvfkstpalfspkiltrslsgrlgn 69
61 CSANESWTEVSWPCRNOSSEKAINIGKKSSWKAFFGVVEKEDSGSTPAKVSAGQRTL 120
70 csanewtevswpcrnossekainigksswkaffgvvekedsgstpakvsagqrtrl 129
121 EYQDSHSGQWRCRLSNVEOCL 141
130 eyqdsghss-gqvlsnveqvl 149

RESULT 6

ABB27961 standard; Peptide: 129 AA.

ABB27961:

01-FEB-2002 (first entry)

Human peptide #612 encoded by breast cell single exon nucleic acid probe.

Human: microarray; single exon probe; gene expression; breast;
disease; cancer.

Homo sapiens.

WO200157271-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00662.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-496933/54.

New spatially-addressable set of single exon nucleic acid probes,
useful for measuring gene expression in sample derived from human
breast, comprises number of single exon nucleic acid probes

Claim 27; SEQ ID NO 10929; 327pp + sequence listing; English.

The invention relates to a spatially-addressable set of single exon
nucleic acid probes for measuring gene expression in a sample derived
from human breast and BT 474 cells. The method involves contacting
the probes with a collection of detectably labelled nucleic acids
derived from mRNA of human breast, and then measuring the label
bound to each probe of the microarray. The probes are useful for
verifying the expression of regions of genomic DNA predicted to
encode proteins. They are useful for gene discovery, and for
determining predisposition and/or prognostic breast disease. Gene

expression analysis is useful for assessing the toxicity of chemical
agents on cells. The microarray of this invention presents a far greater
diversity of probes for measuring gene expression, with far less bias
than expressed sequence tag microarrays. The method is suitable for
rapid production of functional information from genomic sequence. The
present sequence is a peptide encoded by a single exon nucleic acid
probe of the invention.
Note: The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

Sequence 129 AA:

Query Match 52.1%; Score 686; DB 22; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.7e-64;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

16 DDDNTIEFKILAYTRHHVFKSPALFSPKILRTSLSGRLGNCSANESWTEVSWPCR 75
1 dddntiefkilytrhhvfkstpalfspkiltrslsgrlgncsanewtevswpcr 60
76 NSQSEKAINIGKKSSWKAFFGVVEKEDSGSTPAKVSAGQRTLEYQDSHSGQWRCRLS 135
61 nsqsekainigksswkaffgvvekedsgstpakvsagqrtrleyqdsnsgwrcrls 120
136 NVEOCLHE 144
121 nveqclehe 129

RESULT 7

ABB33133 standard; Peptide: 129 AA.

ABB33133:

04-FEB-2002 (first entry)

Peptide #639 encoded by human foetal liver single exon probe.

Human; foetal liver; gene expression; single exon nucleic acid probe.

Homo sapiens.

WO200157277-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00669.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human foetal liver

Claim 27; SEQ ID NO 25768; 639pp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for
measuring human gene expression in a sample derived from human foetal
liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human
 CC fetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 129 AA;

Query Match 52.1%; Score 686; DB 22; Length 129;
 Best Local Similarity 100.0%; Pred. No. 1.7e-64;
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 DDDLTIEFKILAYTRHHVFKSTPALFSPKILRTSLQRLGNCANESWTVEWPCR 75
 Db 1 dddltiefkllaytrhhvfkstpalfspkllrtslsqrglncanestevswpcr 60
 QY 76 NSOSSEKAINLCKKSSWKAFFGVVEKEDSQSTPAKVSAGQRTLEYQDSHSQWSCRLS 135
 Db 61 nsqsekainlgkkskswkaffgvvekedsqstpakvsagqrileyqdsqswscrcls 120
 QY 136 NVEQCLEHE 144
 Db 121 nveqclehe 129

RESULT 8

ABBI8598
 ID ABB18598 standard; Protein; 129 AA.
 AC ABB18598;

DT 23-JAN-2002 (first entry)
 DE Protein #597 encoded by probe for measuring heart cell gene expression.

Human; gene expression; heart; microarray; vascular system;
 cardiovascular disease; hypertension; cardiac arrhythmia;
 congenital heart disease.

OS Homo sapiens.

PN WO200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-48899/53.

Single exon nucleic acid probes for analyzing gene expression in human
 hearts -

Claim 15; SEQ ID No 20368; 530pp; English.

The present invention relates to single exon nucleic acid probes for
 measuring human gene expression in a sample derived from human heart (see
 CC ABA21533-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and proposing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 129 AA;

Query Match 52.1%; Score 686; DB 22; Length 129;
 Best Local Similarity 100.0%; Pred. No. 1.7e-64;
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 DDDLTIEFKILAYTRHHVFKSTPALFSPKILRTSLQRLGNCANESWTVEWPCR 75
 Db 1 dddltiefkllaytrhhvfkstpalfspkllrtslsqrglncanestevswpcr 60
 QY 76 NSOSSEKAINLCKKSSWKAFFGVVEKEDSQSTPAKVSAGQRTLEYQDSHSQWSCRLS 135
 Db 61 nsqsekainlgkkskswkaffgvvekedsqstpakvsagqrileyqdsqswscrcls 120
 QY 136 NVEQCLEHE 144
 Db 121 nveqclehe 129

RESULT 9

AAM53929
 ID AAM53929 standard; Protein; 129 AA.
 AC AAM53929;

DT 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26034.

Human; brain expressed exon; gene expression analysis; probe;
 microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483446/52.

Single exon nucleic acid probes for analyzing gene expression in human
 brains -

Example 4; SEQ ID NO: 26034; 650pp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.

XX Sequence 129 AA:

Query Match 52.1%; Score 686; DB 22; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.7e-64;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 DDDNTIEFKIIAYYTRHHVFKSTPALFSPKLLRTSLRSQRLGNCNSANESWTEVSPCR 75
Db 1 dddntiefkilaytrhhvfksfpalklrltrslrsqrlgncsaneswtevspcr 60
QY 76 NSQSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSNOGORTLEYODSHSQMSRCLS 135
Db 61 nsqsekainlgksswkaffgvvekedsgstpakvsagqrileyqdsnsgwsrcls 120
136 NVEQCLEHE 144
121 nveqclehe 129

RESULT 10
AAM6317
ID AAM6317 standard; Protein: 129 AA.

AC AAM6317;
XX 06-NOV-2001 (first entry)
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 26623.
DE Human bone marrow expressed exon; gene expression analysis; probe;
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukaemia; lymphoma; myeloma.
XX Homo sapiens.

OS WO200157276-A2.
XX 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US00668.
PF 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX Example 4; SEQ ID NO: 26623; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.

XX Sequence 129 AA;

Query Match 52.1%; Score 686; DB 22; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.7e-64;
Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 DDDNTIEFKIIAYYTRHHVFKSTPALFSPKLLRTSLRSQRLGNCNSANESWTEVSPCR 75
Db 1 dddntiefkilaytrhhvfksfpalklrltrslrsqrlgncsaneswtevspcr 60
QY 76 NSQSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSNOGORTLEYODSHSQMSRCLS 135
Db 61 nsqsekainlgksswkaffgvvekedsgstpakvsagqrileyqdsnsgwsrcls 120
136 NVEQCLEHE 144
121 nveqclehe 129

RESULT 11
AAM14186
ID AAM14186 standard; Protein: 129 AA.

AC AAM14186;
XX 12-OCT-2001 (first entry)
XX Peptide #620 encoded by probe for measuring cervical gene expression.
DE Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX Homo sapiens.

XX WO200157278-A2.
XX 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US00670.
PF 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX Claim 27; SEQ ID No 19012; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes
CC (SENP: see A110068-A128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 129 AA;

Db 1 dddlniefkllaylrhvfksfpal fspkllfrtslsqrglgnscaneswvewmpcr 60
 QY 76 NSGSSKAINLGKRSKMAFEGVSEKEDSQSPAPKVSAGQRTLEYODSHSQOWSRCL 135
 Db 61 nsqsssekainlgkkskswkafgvekedsqspakvsagqgrtlleyqdshsqwsrcs 120
 QY 136 NVEQCLEHE 144
 Db 121 nveqclehe 129
 RESULT 14
 AAB85188 standard; protein; 328 AA.
 AC AAB85188;
 XX 07-SEP-2001 (first entry)
 DE Mouse Bcl-G polypeptide.
 XX Bcl-G; cancer; cancer therapy; oncogene; apoptosis; cytostatic;
 KW antiapoptotic; chromosome 12p12.3; mouse.
 XX Mus sp.
 FH Key Location/Qualifiers
 FT Misc-difference 107
 FT /label= unknown
 FT /note= "encoded by NGC"
 XX MO200144282-A2.
 XX 21-JUN-2001.
 PD 13-DEC-2000; 2000MO-US33793.
 PF 14-DEC-1999; 99US-0461641.
 PR (BURN-) BURNHAM INST.
 XX (BURN-) BURNHAM INST.
 PA Reed JC, Godzik A;
 XX WPI: 2001-398125/42.
 DR N-PSDB; AAI22670.
 XX Novel polynucleotide encoding Bcl-G polypeptide, useful for modulating
 XX apoptosis, and for diagnosing and treating cancer
 XX Claim 14; Page 114-115; 117pp; English.
 PS The invention relates to Bcl-G polypeptides and nucleic acids encoding
 XX them. The Bcl-G polypeptides can be expressed by standard recombinant
 CC technology. Bcl-G oligonucleotides (or its anti-sense strand) and Bcl-G
 CC specific antibodies are useful for diagnosing cancer, monitoring cancer
 CC therapy or assessing prognosis of patients with cancer. The Bcl-G
 CC polypeptides are useful for modulating the activity of an oncogene
 CC polypeptides. They are useful for identifying modulators for modulating
 CC a level of apoptosis mediated by the Bcl-G polypeptide, polynucleotide or antibody
 CC composition comprising the Bcl-G polypeptide, polynucleotide or antibody
 CC is useful for treating a pathology characterized by abnormal cell
 CC proliferation especially cancer. The present sequence represents a
 CC mouse Bcl-G polypeptide.
 XX Sequence 328 AA:
 SQ Query Match 38.1%; Score 501; DB 22; Length 328;
 Best Local Similarity 47.5%; Pred. No. 2.5e-44;
 Matches 122; Conservative 33; Mismatches 78; Indels 24; Gaps 8;
 OY 1 MCGTSGCDLEIEIPLDDDLNIEFKILAYTRRHVFKSPALFSPKLLRRLRSLSQKGLGN 60

Db 1 mcsstvydledipreddpnsiefkllaefarhvfknupavfapkslrtslsqkalgt 60
 QY 61 CSANESWTEVSWPCRNSSQSEKAINLCKKSSKMAFEGVSEKEDS-QSTPAKVSAGQCR- 118
 Db 61 ws-tdswtqyslporgspseknlslgkkswwtlfrveakeeglpkspxkelttagppg 119
 QY 119 --TLEYODS-HSQOWSRCLSNVEQCLEHEAVDPKVISIANRVAEIVYSWPPQATQAGCF 175
 Db 120 pfpyergsgfngqhmprslssveqpwrvklywpklyvptewklf----tpghqmsst 175
 QY 176 KSKELFVTEG----LSEQLQGHVPAVSSSKRDEEEOIITAKIVELIKYSGDOLERKDTAFI 231
 Db 176 areeasskrgrfllyffteq---pwdsknkdgedqllskivellkssgddqgreikx-- 230
 QY 232 PIPLYVTSIQGFPDGL 248
 Db 231 ----dkalmssfdqgl 242
 RESULT 15
 AAG65914 standard; protein; 1250 AA.
 AC AAG65914;
 XX 11-FEB-2002 (first entry)
 DE Amino acid sequence of GSK gene id 27142.
 XX Peptide hormone; antidiabetic; anorectic; antianorectic; antiasthmatic;
 KW antidepressant; nootropic; neuroprotectant; hypotensive; hypertensive;
 KW cytostatic; cerebroprotective; vasotropic; human.
 XX Homo sapiens.
 OS WO200172961-A2.
 PN 04-OCT-2001.
 PD 22-MAR-2001; 2001MO-US92226.
 PF 24-MAR-2000; 2000US-192158P.
 PR 28-MAR-2000; 2000US-192668P.
 PR 27-APR-2000; 2000US-200166P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX Agarwal P, Murdoch PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;
 PI Lal Y;
 XX WPI: 2001-639223/73.
 DR N-PSDB; AAI67204.
 XX Isolated polypeptides, which may be peptide hormones, which are
 PT identified by high throughput genome-based biology which identifies
 PT genes and gene products as therapeutic targets for treatment of
 PT diseases such as diabetes and cancer
 XX Claim 1; Page 91-94; 99pp; English.
 PS The invention provides polypeptides (AAG65886-65918) which may be peptide
 CC hormones (including insulin, growth hormones, chemokines, cytokines,
 CC neuropeptides, integrins, kallikreins, lamins, melanins, nutritive
 CC hormones, neuropeptin, pituitary hormones, pleiotrophins, prostaglandins,
 CC secretogranins, selectins, thromboglobulins, thymosins) identified by
 CC high throughput genome-based biology and polynucleotides (AAI67176-67208)
 CC encoding them. The polypeptides can be expressed by standard recombinant
 CC technology. The polypeptides are useful in the treatment of disease such
 CC as diabetes, breast-, prostate-, colon cancer and other malignant tumors,
 CC hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities,
 CC asthma, manic depression, dementia, delirium, mental retardation,

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: June 20, 2002, 02:33:20 ; Search time 2339.76 seconds
(without alignments)
8800.780 Million cell updates/sec

Title: US-09-771-961-1

Perfect score: 984
1 atgtgttagcacccagtggtg.....cacatgaagaagtagactga 984

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenEmbl: *
1: gb_da: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_inv: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query Match Length DB ID Description

1	984	100.0	984	6	AX207692	AX207692 Sequence
2	982.4	99.8	1179	6	AX174783	AX174783 Sequence
3	982.4	99.8	1886	9	AF281254	AF281254 Homo sapi
4	821	83.4	2132	6	AX207696	AX207696 Sequence
5	819.4	83.3	2039	6	AF281255	AF281255 Homo sapi
6	678	68.9	759	6	AX207694	AX207694 Sequence
7	678	68.9	1176	6	AY040274	AY040274 Homo sapi
8	676.4	68.7	954	6	AX174785	AX174785 Sequence
9	505.2	51.3	987	6	AX174823	AX174823 Sequence
10	435.8	44.3	190858	9	AC007537	AC007537 Homo sapi
11	435.8	44.3	194143	2	AC022222	AC022222 Homo sapi
12	435.8	44.3	200499	2	AC091816	AC091816 Homo sapi
13	253.8	25.8	699	9	HSR329005	AJ329005 Homo sapi
14	215.6	21.9	1731	2	AK000127	AK000127 Homo sapi
15	176	17.9	149350	2	AC097248	AC097248 Rattus no
16	175.2	17.8	2877	9	AK026440	AK026440 Homo sapi
17	62.4	6.3	149350	2	AC097248	AC097248 Rattus no
18	42.6	4.3	125020	2	AF429315	AF429315 Homo sapi
19	42.2	4.3	7218	6	166494	166494 Sequence 14
20	39.8	4.0	43	6	AX174816	AX174816 Sequence
21	39.8	4.0	43	6	AX174817	AX174817 Sequence
22	39.6	4.0	180465	9	AC074011	AC074011 Homo sapi
23	39.6	4.0	196769	2	AC092164	AC092164 Homo sapi
24	39.6	4.0	207922	2	AC093164	AC093164 Homo sapi
25	39.4	4.0	125020	9	AF429315	AF429315 Homo sapi
26	39	4.0	63722	2	AC103740	AC103740 Homo sapi
27	39	4.0	176372	2	AC026030	AC026030 Homo sapi
28	39	4.0	177483	9	AC068233	AC068233 Homo sapi
29	37.8	3.8	197852	10	AL060509	AL060509 Mouse DNA
30	37.6	3.8	7055	6	AX006017	AX006017 Sequence
31	37.6	3.8	7783	6	AX006002	AX006002 Sequence
32	37.4	3.8	54632	2	AC094041	AC094041 Rattus no
33	37.4	3.8	71848	2	AC105203	AC105203 Homo sapi
34	37.4	3.8	123218	2	AP004093	AP004093 Oryza sat
35	37.2	3.8	151075	2	AC022594	AC022594 Homo sapi
36	37.2	3.8	166526	2	AC104061	AC104061 Homo sapi
37	37.2	3.8	167132	9	AC069154	AC069154 Homo sapi
38	37	3.8	99102	2	AL356789	AL356789 Human DNA
39	37	3.8	151570	2	AC073923	AC073923 Homo sapi
40	37	3.8	194508	2	AC087473	AC087473 Homo sapi
41	37	3.8	198927	2	AC026863	AC026863 Homo sapi
42	36.8	3.7	645	9	AF270560	AF270560 Homo sapi
43	36.8	3.7	37895	3	U58748	U58748 Caenorhabdi
44	36.8	3.7	267759	2	AC079418	AC079418 Mus muscu
45	36.6	3.7	59857	2	AC109156	AC109156 Mus muscu

ALIGNMENTS

RESULT 1
AX207692
LOCUS AX207692 984 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 1 from Patent WO0157213.
ACCESSION AX207692
VERSION AX207692.1 GI:15422374
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 984)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS

Donoho, G., Hilbun, E., Turner, C.A., Friedrich, G.B., Abulin, A.,
Zambrowicz, B. and Sands, A.T.
Human bcl-x-like proteins and polynucleotides encoding the same
Patent: WO 0157213-A 1 09-AUG-2001;
Lexicon Genetics Incorporated (US)

FEATURES

source

1. 984
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 289 a 230 c 250 g 215 t
ORIGIN


```

QY 901 ctgaagaagaactctcgcacatgatccagcagcgttgatgggaaaaatacttggg 960
DB 1096 CTGAAGAAGAAGACTTCTCGCATGATCCAGACGACGCGTGTGAGGAAAAAATCTTGGG 1155
QY 961 atatacatgaagaagttaactga 984
DB 1156 ATATCATCATGAAGAAGTAGACTGA 1179

RESULT 4
LOCUS AX207696 2132 bp DNA linear PAT 31-AUG-2001
DEFINITION Sequence 5 from Patent WO0157213.
ACCESSION AX207696
VERSION AX207696.1 GI:15422376
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2132)
Donoho, G., Hilbun, E., Turner, C.A., Friedrich, G.B., Abuin, A.,
Zambrowicz, B. and Sands, A.T.
Human bcl-x-like proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0157213-A 5 09-AUG-2001;
Lexicon Genetics Incorporated (US)
FEATURES
SOURCE 1.2132
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 615 a 477 c 486 g 549 t 5 others
ORIGIN
Query Match 83.4%; Score 821; DB 6; Length 2132;
Best Local Similarity 86.5%; Pred. No. 3,4e-224;
Matches 984; Conservative 0; Mismatches 0; Indels 153; Gaps 1;

QY 1 atgtgtgagcagcagtggtgtgacgaagaatccccatagatgatgaacaaac 60
DB 192 ATGTGTGAGCAGCAGTGGGTGTGACCTGGAGAAATCCCCATGATGATGATCACTAAGC 251
QY 61 accatagaattcaaaatcctgcctactacacacacacacacacacacacacac 120
DB 252 ACCATGAGAAATTCCTGCTGCTACTACACACACATCATGCTTCAAGACACCCCT 311
QY 121 gctctcttcacaaagcgtgtgagacaaagaattgtccacagagggccttgggaat 180
DB 312 GCTCTCTTCACCAAGCTGCTGAGAACAAAGATTGTCCACAGAGGGCCTTGGGGAAT 371
QY 181 tgttcagcaaatagtcataagcagaggtgtcatggtccttcagagaattcccaatcagt 240
DB 372 TGTTCGCAAAATAGTCATGACAGAGAGTGTCAATGCTTGCAGAAATTCCTCAATCCACT 431
QY 241 gagaagcagcataaaaccttggcagaanaaagtctcttggaaagcattcttggagtagtg 300
DB 432 GAGAAGCCATAAAGCTTGGCAAGAAAAAGCTTCTTGAAGACATCTTTTGGAGTAGTG 491
QY 301 gagaaggaagatcgcagagacagccgcgcgaaggtctgtcgaaggtcgaaggaagtg 360
DB 492 GAGAAGGAAGATTTCGACAGACACGCTGCCTCAAGGCTCTTGTCTCAGGGTCAAAAGAGCTTG 551
QY 361 gaatacaagatcgcagcagcagcagtggtccaggttcttccaaagtgagcagtg 420
DB 552 GAATACCAAGATTCGACACACAGCAGCAGTGTCTTTTCAACGTGAGCAGCTGC 611
QY 421 ttggagcagtagcgtgtgagcccaaaagtcattccatctccaacagtagcagtaatt 480
DB 612 TTGGACATGAAGACTGTGGAGCCCAAAAGTCAATTCATTCGCCAACCGAGTAGCTGAATTT 671
QY 481 gtttatctctgacacacacacagcagcagcagcaggaaggttcaagtccaagaagatt 540

```

```

DB 672 GTTTATTCCTGGCCACACACACACAGCCAGGAGGAGGCTTCAATGCCAAGAGATT 731
QY 541 ttgttaactgagaggtctctccttcacagctcgaagcagctgtcgttagctcaactctc 600
DB 732 TTTGTACTGAGGGGTCTCTCTTCAGCTCCAGGCGACAGTGGCTTGTAGTTCAAGTTCT 791
QY 601 aagaagaatgaagaagaacaaatactagcacaataattgttgagctgtcgaataatcaga 660
DB 792 AAGAAGATGAAGAAGAACAAATACTAGCCAAAATGTGAGCTGCTGAATTTACAGGA 851
QY 661 gatcaattggaagaaga----- 676
DB 852 GATCAGTTGGAAAGAAAGAGACACTGCTTCATCCCATTCCTTGGTTGACACACGATC 911
QY 677 ----- 676
DB 912 CAGGGTTTCCACAGGATGTTTGTATGGCTGCATTTAGCTAAAGATGAAGTCTGTCTC 971
QY 677 ----- agctgaagaa 687
DB 972 TGCTCTGGAGCCAGCTACTGATGAGTCTTATTTTGTACACAGCTGAAGAAA 1031
QY 688 gataagctttgatgggacactccagagtaggctgtcctactctgttttcaagaccatc 747
DB 1032 GATAAGGCTTTGATGGGCGACTTCCAGGATGGGCTGTCTTCTTTTCAAGACCATC 1091
QY 748 acaagaccagctcctaattgtgtgagcccaagggagaaatcagaaggtcgaagctcagagc 807
DB 1092 ACAGACCCAGTCTTATATGGTGTGACCCAGGGAGGAGATCAAGGTCAGAGGC 1151
QY 808 tttaagctgccttgtatagagctcaagcgaagctcaagctatttgaaccaccgc 867
DB 1152 TTTAAGGCTGCCCTTGTATATGACGTACAGGCCAGACCTCAACCTATTGCAACACCCG 1211
QY 868 atgaacaaggtctcgtgagcttggcaccagtaactcgaagaagaactctcgcacatgac 927
DB 1212 ATGAACAGGCTCTGGCTTTGGACCAAGTACCTGAAGAGACATTTCTGCGATGATC 1271
QY 928 cagcagcagcttgatgggaaaaataacttggatatacatgaagaagttagactga 984
DB 1272 CAGCAGCAGCGTGTGATGGGAAAAAATACTTGGGATTCATCATGAGAGTAGACTGA 1328

RESULT 5
AF281255 2039 bp mRNA linear PRI 28-JAN-2001
LOCUS AF281255
DEFINITION Homo sapiens apoptosis regulator BCL-6 short form (BCL6) mRNA,
complete cds.
ACCESSION AF281255
VERSION AF281255.1 GI:12584086
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2039)
AUTHORS Guo, B., Godzik, A. and Reed, J.C.
TITLE Bcl-6, a novel pro-apoptotic member of the bcl-2 family
JOURNAL J. Biol. Chem. 276 (4), 2780-2785 (2001)
MEDLINE 21264734
REFERENCE 2 (bases 1 to 2039)
AUTHORS Guo, B., Godzik, A. and Reed, J.C.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-2000) The Burnham Institute, 10901 North Torrey
Pines Road, La Jolla, CA 92037, USA
FEATURES
SOURCE Location/Qualifiers
1..2039
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
1..2039
/gene="BCL6"
196..954
CDS

```

/gene="BCLG"
/note="member of Bcl-2 family"
/codon_start=1
/product="apoptosis regulator BCL-G short form"
/protein_id="AAG59794.1"
/db_xref="GI:12584087"
/translation="MSTSGCDLEIPLDLDDDLNTEIKILAYVYRHHVFKSTPALES
PKLIRTSLSORGLGNCSEKSWTSPKRNSSSEKAINLKKSKMKAFFGVK
EDSSQTPAKVSAQGRITLEYODSHSQMRSRCSNCEHEHVAIDPKEVISINARVAE
VSWPPOATQAGFSKSEIFVTEBGLSFQLOGHVAVASSSKDEEQLIAKIVELLKY
SGDLEKRDNAFIRIPLYDTSIOGPPDGLMACI"

BASE COUNT 567 a 475 c 484 g 513 t

ORIGIN

Query Match 83.3%; Score 819.4; DB 9; Length 2039;
Best Local Similarity 86.5%; Pred. No. 9,7e-224;
Matches 983; Conservative 0; Mismatches 1; Indels 153; Gaps 1;

1 atgtgttagcaccagtggtgtgtgacctggaagaatacccccctagatgatgatacctaac 60
196 ATGTGTAGCACACAGTGGGTGTGACCTGGAGAATAATCCCCCTGATGATGATGACCTTAAC 255
61 accataagaattcaaatctctgctactactaacacagacatcatgcttcaagagcaccct 120
256 ACCATAGAATTCAAATCTCTGCTACTACACACAGACATCATGCTTCAAGACACACCCCT 315
121 gctctctctaccaaagctctgtgaaacaaagtgttcccaagagggcctgagggaat 180
316 GCTCTCTTCTTCCAAAGCTGCTGAGAACAGAGATTGTCTCCAGAGGGGCTTGGGGAAT 375
181 tgttcagcaaatgtatcgtacagagaggtgtcatgtgctctgagaataatcccaatcaat 240
376 TGTTCAGCAAAATGTATGTATGACAGAGGTGTATGCTTGGAGAAATTCACATCCAGT 435
241 gagaagagccataaacctgtgcaagaagaagtcttctggaagaatcttctggaatgtg 300
436 GAGAAGAGCCATAAACCTTGGCAAGAAAAAGTCTTCTGGAAGACATCTTGGAGTGTG 495
301 gagaagagaagtgtgcaagagcagcgtcccaaggtctgtctgcaagggttcaagaagcttg 360
496 GAGAAGAGAAATTTGGCAGACAGACGCTCCAGAGGTCTGTCTCAGGGTCAAGAGACGTTG 555
361 gaatacagaattcgcacagcagcagcagtggtccaggtgtcttcttaactgagcagtg 420
556 GAATACCAAGATTGCGACAGCAGCAGAGGTGTCTCAGGTGTCTTCTTAACGAGCAGATGC 615
421 ttggagcattgagctgtgagcccaagtcatcttccattgccaacccagtagtgaatt 480
616 TTGGAGCATGAAAGCTGTGGACCCCAAGTCATTTCCATTGCCAACCAGATGAGCTGAAT 675
481 gttattcccttgccacccacccaagcgaagcgaagagaggtcttaagtccaagaagt 540
676 GTTATTCTCTGGCCACACACACAGCAGCAGGAGGAGGCTTCAAGTCCAAAGAGAT 735
541 ttgttaactgaggtgtctctcttccagctccaagcagcagtgctgttagcttaagtct 600
736 TTGTTAAGTGAAGGCTCTCTCTCCAGCTCCAAAGGCAAGGCTGCTGTAGTTCAAGTCT 795
601 aagaagaatgagaagaacaatactactagccaataatgttgaagctgtgtaattcagga 660
796 AAGAAGAATGAGAAGAACAATACTAGCCAAATGTGTGAGCTGTGAATATTTACAGA 855
661 gatcagttggaagaaga----- 676
856 GATCAGTTGGAAAGAGAGACACTGCTTCATCCCATTCCTTGTGTTGACACACGATC 915
677 ----- 676
916 CAGGGTTTTCACAGAGATGTTGATGGCTTGATTTGAGCTAAAGATGACCTTCTGTC 975
677 ----- agtgaagaaga 687
677 -----

Db 976 TGCTCTGTGGAGCCCAAGCTACTGTACTGAGTGTCTTATTTCTTTGTACACAGCTGAAGAAA 1035
Qy 688 gataagctttgtagggccactcccaagatgagctgtctactactgttttcaagaacatc 747
Db 1036 GATAAGGCTTTGATGGGCCCTTCCAGAGATGGGCTGTCTCTACTCTGTTTCAAGCCATC 1095
Qy 748 acagaccaggtcttaattggtgtgtgaccccaaggaggagaatacagaagttcaaaagctcaagg 807
Db 1096 ACAGACCAGAGTCTTATATGGGTGTGACCCCAAGGAGGAATCAGAGGTCAAGGCTCAGGGC 1155
Qy 808 ttaagctgcgcctctgtaataagacgltacagcccaagctcaagctatattgaacaccaccg 867
Db 1156 TTTAAGGCTGCCCTTCTTAATGAGCTCACGCCCAAGCTCACAGCTATTGACCAACCCACCG 1215
Qy 868 atgaacaggtctctggtgcttggtgacccaagtlaccctgaagaagactctgcgcattgac 927
Db 1216 ATGACACAGGCTCTGGGCTTTGGCACCAGTACTGTAAGGAACCTTCCGCAATGAGT 1275
Qy 928 cagcagcagcgtgtgagtgaggaataaatacttggagatatacagaagaagttagactga 984
Db 1276 CAGCAGCAGGCTGATGGGAAAAAATCTTGGGATATCATGAAAGATGAGACTGA 1332
RESULT 6
AX207694 759 bp DNA linear PAT 31-AUG-2001
LOCUS AX207694
DEFINITION Sequence 3 from Patent WO0157213.
ACCESSION AX207694
VERSION AX207694.1 GI:15422375
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 759)
AUTHORS Donoho, G., Hilbun, E., Turner, C. A., Friedrich, G. B., Aduin, A.,
Zambrowicz, B. and Sands, A. T.
TITLE Human bcl-x-like proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0157213-A 3 09-AUG-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
1..759
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 216 a 182 c 185 g 176 t
ORIGIN

Query Match 68.9%; Score 678; DB 6; Length 759;
Best Local Similarity 100.0%; Pred. No. 3.1e-183;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgtgttagcaccagtggtgtgtgacctggaagaatacccccctagatgatgatacctaac 60
Db 1 ATGTGTAGCACACAGTGGGTGTGACCTGGAGAATAATCCCCCTGATGATGATGACCTTAAC 60
Qy 61 accataagaattcaaatctctgctactactaacacagacatcatgcttcaagagcaccct 120
Db 61 ACCATAGAATTCAAATCTCTGCTACTACACACAGACATCATGCTTCAAGAGCACCCCT 120
Qy 121 gctctctctaccaaagctctgtgagacaagaagtgttcccaagagggcctgagggaat 180
Db 121 GCTCTCTTCTTCCAAAGCTGTGAGAAACAAGAAGTTGTGCCAGAGGGGCTGGGGAAT 180
Qy 181 tgttcagcaaatgtatcgtacagagaggtgtcatgtgctctgagaataatcccaatcaagt 240
Db 181 TGTTCAGCAAAATGTATGTATGACAGAGGTGTCTCAGGCTTCCAGAAATTCACATCCAGT 240
Qy 241 gagaagagccataaaccttggcagaagaagaatctctcttggagaagacttctggagtagtg 300
Db 241 GAGAAGGCCATTAACCTTGGCAGAGAAAGATCTTCTTGGAAAGCATTTTGGAGTATG 300
Qy 301 gagaagagaattcgcagagagcagcctgccaaggtctctgtcaggtcagaagaagcgttg 360

Df	301	GAGAAAGGAAATTTGGCAAGACACGCCCTGCCAAGGTCTCTGCCTCAGAGTCCAAAGACAGTTC	360
Oy	361	gaataccaagttggcaaacgacagaatggtccaaagtcttcttaacagtggaagatgc	420
Df	361	GAATPACCAAGATTGGCGACACAGCGACGATGGTGCCAGGTCTTTTAACGTGAGCATGC	420
Oy	421	ttagagcatgaagctgtgtgacccccaaaattcatttcattcgacaacaggatagctaat	480
Df	421	TTCGAGCATGAAGCTGTGTGACCCCCAAATTCATTTCATTGSCAACCGACTAGCTAAATT	480
Oy	481	gttatctccgtgccaccaaccaagaagcgaccacagcaggacttcaagtlccaagaatt	540
Df	481	GTTTATTCTTCGGCCACCACCAAGCGACCCAGGAGAGGCTTCAATGCCAAAAGATT	540
Oy	541	tttctaactggaggtctctctctccagttccaaagcacgctgctgtatgcttaagttct	600
Df	541	TTTTGTAATCTGAGGGTCTCTCTCTCCAGCTCCAAAGGCCACGTGCTGTACAAGTTCT	600
Oy	601	aagaagaatlgaagaagaacaatactagccaanaattgttgaagctgtcgtgaatatcca	660
Df	601	AAGAAGATGAAAGAAAGAACAAATACTACCCAANAATTGTTGAGCTGCTGAATAATTCACGA	660
Oy	661	gatcagttggaagaagaag 678	
Df	661	GATCAGTTGGAAAGAAAG 678	
RESULT	7		
LOCUS	AY040274	1176 bp	mRNA linear PRI 01-AUG-2001
DEFINITION	Homo sapiens apoptosis regulator BCL-G median form (BCLG) mRNA,		
ACCESSION	AY040274		
VERSION	AY040274.1	GI:15072494	
KEYWORDS	human,		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1176)		
TITLE	Montpetit,A., Boily,G. and Sinnett,D.		
JOURNAL	A detailed transcriptional map of the chromosome 12p12 tumor		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 1176)		
TITLE	Montpetit,A. and Sinnett,D.		
JOURNAL	Direct Submision		
REFERENCE	Submitted (14-JUN-2001) Hemato-Oncology, Hopital Ste-Justine, 3175		
AUTHORS	Cote-St-Catherine, Montreal, QC H3T 1C5, Canada		
TITLE	Location/Qualifiers		
FEATURES	1..1176		
source.	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="12"		
	/map="12p12"		
	1..1176		
	/gene="BCLG"		
	8..838		
gene	/gene="BCLG"		
	/note="member of the BCL-2 family"		
	/codon_start=1		
	/product="apoptosis regulator BCL-G median form"		
	/protein_id="AAK22109.1"		
	/db_xref="GI:15072495"		
CDS	/translation="WCSTSGCDLEELPLDDDDINTIEFKILAYTRHHVKSPPALFSS		
	PKLIRTSLSORGLGNCSEANEWSWTEVSWPCRNSSSRKAINLGRKKSMRAFGVEEK		
	EDQSTRFAKYSAOGARTLEYODSHSQOMRSLSNVEQCLEHNAVDPRVISIANRVARI		
	VYMWPRQAOAGAGFKSKETFEVGESFOLOGHPVASSSSKKDDEEOIILAKIYEILKY		
	SGQLERKEKEWESESATRTIKMLIMRSSGGPRVLISIPAPFNKNSTLTLPPESEHN		
BASE COUNT	363 a	277 c	270 g 266 t
ORIGIN			

Query Match	Similarity	Score	DB	Length
Best Local	98.6%	Pred. No. 3.2e-183		
Matches	684	Conservative	0	Mismatches 10; Indels 0; Gaps 0;

Query	1	atgtltagacacagatggtgtgtltagaccttggaagaatccccctagatgatgattacctaac	60
Db	8	atgttagacacagatggtgtgtltagaccttggaagaatccccctagatgatgattacctaac	67
Qy	61	accataagaattcaaaatcctctgctactaacacagacaatcatatgttcttaagaagaccctc	120
Db	68	accataagaattcaaaatcctctgctactaacacagacaatcatatgttcttaagaagaccctc	127
Qy	121	gtctctctccacaaagctgtctggaacaaagattgttcccaagaggcccttgaggat	180
Db	128	gctctctctccacaaagctgtctggaacaaagattgttcccaagaggcccttgaggat	187
Qy	161	tgttcaagaatagtcatactgacagaagtgatcatgagccttgagaattcccaatcagc	240
Db	188	tgttcaagaatagtcatactgacagaagtgatcatgagccttgagaattcccaatcagc	247
Qy	241	gagaagggccataaactctggcgaagaagaagcttcttggaaagcattcttggagtagt	300
Db	248	gagaagggccataaactctggcgaagaagaagcttcttggaaagcattcttggagtagt	307
Qy	301	gagaaggaagaattcgcagaagcagcgccttcgaagctctgctcgaaggtcgaagaagcgttg	360
Db	308	gagaaggaagaattcgcagaagcagcgccttcgaagctctgctcgaaggtcgaagaagcgttg	367
Qy	361	gaataccaaagattgcacaaagcgaagcgaagtgatcgaagtgatccttctaagctggagcagtc	420
Db	368	gaatccaaagattgcacaaagcgaagcgaagtgatcgaagtgatccttctaagctggagcagtc	427
Qy	421	ttggaagcattgaagctgtgtgaccccaagctatcttcatcttgccaacgagtagctgaatt	480
Db	428	ttggaagcattgaagctgtgtgaccccaagctatcttcatcttgccaacgagtagctgaatt	487
Qy	481	gttattccttgccaccacacaaagcgaagcgaagcgaagcgttcaaatgccaaagaagat	540
Db	488	gttattccttgccaccacacaaagcgaagcgaagcgaagcgttcaaatgccaaagaagat	547
Qy	541	tttgaacttgaggtgtcttctcttccagctccaagccagcgtgagctgattccaagtct	600
Db	548	tttgaacttgaggtgtcttctcttccagctccaagccagcgtgagctgattccaagtct	607
Qy	601	aagaaagaatgaaagaacaaatactagccaagaattgtgtgagctgtcgaataatcaga	660
Db	608	aagaaagaatgaaagaacaaatactagccaagaattgtgtgagctgtcgaataatcaga	667
Qy	661	gatcagcttggaagaagaagctgaagaagaataagg	694
Db	668	gatcagcttggaagaagaagctgaagaagaataagg	701

LOCUS	AX174785	954 bp	DNA	linear	PAT 03-JUL-2001
DEFINITION	Sequence 3 from Patent WO0144282.				
ACCESSION	AX174785				
VERSION	AX174785.1	GI:14598301			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 954)				
JOURNAL	Reed, J.C. and Godzik, A.				
FEATURES	Bcl-9 polypeptides, encoding nucleic acids and methods of use				
SOURCE	Patent: WO 0144282-A 3 21-JUN-2001;				
	The Burnham Institute (US)				
	Location/Qualifiers				
	1..954				

/organism="Homo sapiens"
/db_xref="taxon:9606"
196..954
/note="unnamed protein product"

CDS

/protein_id="CAC43801.1"
/db_xref="GI:14598302"
/translation="MCSISGCDLEELPLDDDLNTEIEKILAYVTRHHVFKSPALFS
PILRTSLRSORGLGNCNANESWTEVMSRNSQSEKALNIGKSSMKAFGVVEK
EDSOTPAKVSAGORTLEODSHSDMSRCLNVEQCLEHVAIDPKYISIANRVAEI
VYSPMPOTAGGFKSKETFEVNEGLSFOLOGHVPVASSSKKDEEQLIAIVELKY
SSDQERKDTAFPIPLVDTISIGFPDGMACI"

BASE COUNT 267 a 237 c 240 g 210 t
ORIGIN

Query Match 68.7%; Score 676.4; DB 6; Length 954;
Best Local Similarity 99.9%; Pred. No. 9.2e-183;

Matches 677; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 atgtgtgaagaccagtggtgtgtgaacctggaagaatccccctagatgatgatacctaac 60
196 ATGTGTGAGACACCACTGGGTGTGACCTGGAGAAATCCCCCTAGATGATGATCCTTAAC 255
61 accatagaattcaaaaacccctgactactacacacacacacacacacacacacacac 120
256 ACCATVGAATTCAAAATCCCTGGCTTACTACACACACATCATGCTTCAAGAGCACCCCT 315
121 gctctcttaccacaaagctgctgagaaagaattgtctccagaaggcctgaggaaat 180
316 GCTCTCTTCTCACCACCAAGCTGTGAGAACAGAGTTGTCCACAGAGGGCTGGGGAAT 375
181 ttttcagcaaatagatcatgagaaaggtgtcatgaccttgagaagaattcccaatccagt 240
376 TTTTCAGCAAAATAGTCATGACAGAGAGTGTGATGCTTGCAGAAATTCCTCAATCCAA 435
241 gagaagagcacaataaacttggcagaagaagaagctctctcttggaaagcaatcttggagtagtg 300
436 GAGAAGGCCATTAACCTTGSCAAGAAAAGCTCTTGGAAGCATCTCTTGGAGTAGTG 495
301 gagaagaagaattcgcagaaacagcctgccaaggtctctgtcagaaggtcacaagaagacgttg 360
496 GAGAAGGAAGATTGCGAGACAGCCTGCAAGGCTCTGCTCAGGGGTCAAGGAGACGTGG 555
361 gaatacgaagattcgcacacacagcagcagtggtcagagtgcttcttaagctggagacgttcg 420
556 GAATACCAAGATTTCGACACACACAGCACTGTGTCAGGTCTTTCTTACGTGAGACAGTGC 615
421 ttggagacatgaagctgttggaccccaagtaattccattgccaacccagtagctgaatt 480
616 TTGGACCATGAAGCTGTGAGACCCCAAGTATTTTCATTGGCAACCGAGTAGCTGAATTT 675
481 gtttattctctggcacaac 540
676 GTTTACTCTCGGCGACACACACACAGCAGCAGGAGGAGGCTTCAAGTCAAGAGATT 735
541 ttgtgaactgaggtgtctctctctcagctcccaagcagcagctgagcttcaagttct 600
736 TTTGTAACTAGAGGCT 795
601 aagaagaatgaagaagaacaataactagccaataatgttgaagctgtcgaataatcaga 660
796 AAGAAAGATGAAGAAGAAATCTAGCCAAATTTGTTGAGCTGCTGAATATTTCAGGA 855
661 gatcagttgagaagaagaag 678
856 GATCAGTTGGAAAGAAAG 873

RESULT 9
AX174823 AX174823 987 bp DNA linear PAT 03-JUL-2001
LOCUS
DEFINITION Sequence 41 from Patent WO0144282.

ACCESSION AX174823
VERSION AX174823.1 GI:14598319
KEYWORDS
SOURCE
ORGANISM

house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 987)
AUTHORS
Reed, J.C., and Godzik, A.
TITLE
Bel-g polypeptides, encoding nucleic acids and methods of use
JOURNAL
Patent: WO 0144282-A 41 21-JUN-2001;
The Burnham Institute (US)

FEATURES
source
Location/Qualifiers
1..987
/organism="Mus musculus"
/db_xref="taxon:10090"

CDS
1..987
/note="unnamed protein product"

/codon_start=1
/protein_id="CAC43827.1"
/db_xref="GI:14598320"
/translation="MCSISVYDLIEDPLEDDDNSTIEKILAFYARHHVKNPAYS
PKLSRTSLSKALGTSWTDSTVOYSLRCRSPSSKRNISLKKSSMTLPVAAKE
EGLPXSPKEIRAGQGPFPVERQSGFNHMPRLSVQEPKRVLMIPKWLVPTE
WLKLPFGHQMSTAREBASRGRRLYRFGPMDSKKDEBDIISKIVELIK
SSFDGLREIKDKRALMSFODGLSTFKITDILFDLRDGESEVARGFARFALA
IDATAKITASDNHNPMLRGCTKYLKEYFSPWVOONGEMKILGISHEVD"

BASE COUNT 272 a 246 c 271 g 197 t 1 others

Query Match 51.3%; Score 505.2; DB 6; Length 987;
Best Local Similarity 72.9%; Pred. No. 1.1e-133;

Matches 728; Conservative 0; Mismatches 244; Indels 27; Gaps 5;

1 atgtgtgaagaccagtggtgtgtgaacctggaagaatccccctagatgatgatacctaac 60
1 ATGTGTGAGACACCACTGGGTGTGACCTGGAGAAATCCCCCTAGATGATGATCCTTAAC 60
61 accatagaattcaaaaacccctgactactacacacacacacacacacacacacacacac 120
61 AGCATAGAGTTCAAAATCCCTGGCTTACTACGCCAGACCATGTCTTCAAGAAACCCCG 120
121 gctctcttaccacaaagctgctgagaaagaaggtgttccagaaggcctgaggaaat 180
121 GCTGCTTCTTCCGCCAAGCTTCCAGAACAGAGTGTGCTGCCAGAAAGCCCTGGGAGCT 180
181 ttgtcagaatagatcatgagaaaggtgtcaatgagccttgcagaatcccaatccagt 240
181 TGTG---CAACTGATTCCTCGACACAGGATATCTGCTGTCAGAGGTTCCCTCCAGC 237
241 gagaagcacaataaacttggcagaagaagaagctctcttggaaagcaatcttggagtagtg 300
238 GAAAAGAACATCAGTGTGGGCAAGAAAGATCTTCTTGGAACACCTCTCAAGGAGGCC 297
301 gaga---aggaagaattcgcagacagcagcagcagcagcagcagcagcagcagcagcag 349
298 GAGAAGAGAGAGGCGCTGCGGAGCTGCCCAAGAGAAATCCGAGCTCAGAGGCTCCAGAGGC 357
350 ---aaagacgttggaaatcacaagaattcgcacagcagcagcagcagcagcagcagcagc 405
358 CCCTTCCCGGAGACCGGAGAGTGGCTTCCACAAACAGCAGCAGCTGCGCCAGAGTCTAGC 417
406 aagctgagagcagtggttggagacatgaagcttggagcccaagaatcattcatatggcaac 465
418 AGTGTGAGAGAG-CCCTGGAGAGTGAAGTTGTGATTCCAAAGTGGCTTATTGGCCAAC 476
466 cgaatagctgaatatttattctcttgcacacacacacacacacacacacacacacacacac 525
477 AGAGTGGCTGAATGTTTACTCTGCGCACACACAGATGTCATCCACACAGGAGGAGA 536
526 aagtcacaagaagatttctgaactgaggtgtctctctccagctcccaagcagcagtgctc 585

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	SOURCE
Db 537	ACCCACTCAAGAGAGAGGTCCTCGAGATTTTGTACTTACCTCAGGTTGAAAGACCTTGGA	596				
Qy 586	gtaagctcaagcttcaagaagaatgaaagaacaaataactctgccaataatgtgtgaactg	645				
Db 597	CT-----CTAAGAATTAAGATGCTGAAGACCAAAATATATAAGCAAGATGTGGACCTG	648				
Qy 646	ctgaatatcagaagatacaattggaagaagaagctggaagaagaataagcttgatggc	705				
Db 649	CTGAATTCCTCGGGGCACTTGGGAAGAGAGATTAAGCAAGACAGCTTTGATGACG	708				
Qy 706	caattccaggatgagctgtctcactctgttttcaagaccatacagaaccagtcctaattg	765				
Db 709	AGCTTCACAGACGCGGCTGTCTTACTCAACGTTCAAGACCATCACAGACCTGTTCCTGAGG	768				
Qy 766	gtgtgtgagcccccaggaggagaatcagagttcaagttcaagctcaggcttctaagctgccttgta	825				
Db 769	GACGTGGACACCCAGAGAGAAATCAAGAGGTCAAAAGCTCGGGGCTTCAAGGCTGCCCTTGCA	828				
Qy 826	atagaagtcacagcccaagctcacagctattgacaacaccagccgaatgaacaggtccctggc	885				
Db 829	ATAGAGCCCATTCGCCAAGCTCACGGCATCGGACACACACCACATGATAGAACTGTGGC	888				
Qy 886	tttgcacccaagtaacctgaaagaagaactctctgcatactgataccagcagcaggtgtgattg	945				
Db 889	TTTCGGAGCCAAATACCTTAAGAGACTTCTCCCTCCCTGGGTTGACAGAGATGGGGATGG	948				
Qy 946	gaaaaaatactgttgatatacatcatgaaagaagtactga	984				
Db 949	GAAAAAATACCTTGGGATCTCATATGAAGATGACCTGA	987				
RESULT 10						
AC007537	AC007537	190858 bp	DNA	linear	PRI 09-OCT-1999	
LOCUS	AC007537	190858 bp	DNA	linear	PRI 09-OCT-1999	
DEFINITION	Homo sapiens 12p12 BAC RPi11.267J23 (Roswell Park Cancer Institute					
ACCESSION	AC007537					
VERSION	AC007537.3	GI:4914348				
KEYWORDS	HTG.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
1 (bases 1 to 190858)						
Muzny,D., Arenson,A.D., Bouck,J., Bunac,C., Chen,Z., Ding,Y.,						
Dugan,S., Durbin,U., Forcum,U., Garcia,C., Gorrell,J.H.,						
Gorrell,L.B., Hernandez,J., Issar,A., Jackson,L., Kneitz,S.,						
Kondejewski,N., Lau,S., Leal,B., Lee,E., Lichtarge,O., Liu,W.,						
Logan,O., Lu,J., Marondel,I., Martinez,C., Merscher,S., Miller,A.,						
Montgomery,K., Oswal,G., Pampell,L.R., Parish,B.J., Perez,L.,						
Rashid,N.D., Rives,C., Scherer,S.E., Shen,H., Shim,C., Simon,M.,						
Vo,O., Williamson,A., Worley,K.C., Xiang,A.M., Yang,R., Yu,W.,						
Zhou,X., Kucherlapati,R., Nelson,D. and Gibbs,R.A.						
Direct Submission						
Unpublished						
2 (bases 1 to 190858)						
Worley,K.C.						
Direct Submission						
Submitted (13-MAY-1999)	Molecular and Human Genetics, Baylor					
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA						
3 (bases 1 to 190858)						
Worley,K.C.						
Direct Submission						
Submitted (27-MAY-1999)	Human Genome Sequencing Center, Department					
of Molecular and Human Genetics, Baylor College of Medicine, One						
Baylor Plaza, Houston, TX 77030, USA						
4 (bases 1 to 190858)						
Worley,K.C.						
Direct Submission						
Submitted (28-MAY-1999)	Human Genome Sequencing Center, Department					
of Molecular and Human Genetics, Baylor College of Medicine, One						
Baylor Plaza, Houston, TX 77030, USA						
5 (bases 1 to 190858)						
REFERENCE						
TITLE						
AUTHORS						
JOURNAL						
REFERENCE						
TITLE						
AUTHORS						
JOURNAL						
REFERENCE						
TITLE						

AUTHORS
TITLE
JOURNAL

Worley, K.C.
Direct Submission
Submitted (09-OCT-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 28, 1999 this sequence version replaced gi:4895156.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
sc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats were identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 23:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8058/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

Summary Statistics	
Contig length:	190858
Phrap values in estimate:	188824
Average error rate (BCM-Phrap estimate):	0.000239024
Fraction of Phrap values less than 40 :	0.039746
Number of consensus changing edits:	12
Number of N's in consensus :	0

Position	Consensus	Changing	edits	Edited+Context
16186	atatgttga(a)(c)cccatctct	atattgttga(a)(a)cccatctct		atattgttga(a)(a)cccatctct
26335	taaatattaca(m)acggacaana	taaatattaca(a)acggacaana		taaatattaca(a)acggacaana
35016	ttttttttt(m)aggcgagctc	ttttttttt(g)aggcgagctc		ttttttttt(g)aggcgagctc
35923	ttatatatt(m)ttttttttt	ttatatatt(t)ttttttttt		ttatatatt(t)ttttttttt
49346	tgccagcgcgga(m)cttgcagatga	tgccagcgcgga(g)cttgcagatga		tgccagcgcgga(g)cttgcagatga
49643	tgccacacat(m)gttaaaaccc	tgccacacat(a)gttaaaaccc		tgccacacat(a)gttaaaaccc
80511	ggagccggag(m)tgacaacagag	ggagccggag(c)tgacaacagag		ggagccggag(c)tgacaacagag
91836	ggaagaaga(m)acacaccccttg	ggaagaaga(a)aacacaccccttg		ggaagaaga(a)aacacaccccttg
104278	accacattaa(g)(t)tagtagacag	accacattaa(g)ttagtagacag		accacattaa(g)ttagtagacag
114732	ctccctccct(m)ctccctccctc	ctccctccct(c)ctccctccctc		ctccctccct(c)ctccctccctc
185043	catgtgtgaa(m)cccgctctcta	catgtgtgaa(c)cccgctctcta		catgtgtgaa(c)cccgctctcta
189415	gattcatlta(m)cccagaagac	gattcatlta(a)cccagaagac		gattcatlta(a)cccagaagac

----- Distribution of Quality < 40 Bases -----

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

2 (bases 1 to 194143)
Unpublished
Direct Submission
Submitted (27-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2001 this sequence version replaced g1:13957543.

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: HABL
Center clone name: RP11-52513
Summary Statistics
Sequencing vector: Plasmid: M77789
Sequencing vector: M13: L08821
Chemistry: Dye-terminator Big Dye 3 of reads
Assembly program: Phrap version 0.990329
Consensus quality: 185441 bases at least Q40
Consensus quality: 200072 bases at least Q30
Consensus quality: 208935 bases at least Q20
Estimated insert size: 199196: sum-of-coverage estimation
Quality coverage: 0x in Q20 bases; agrose-gel estimation
Quality coverage: 3.3x in Q20 bases; sum-of-coverage estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 26 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 21476: contig of 21476 bp in length
21477 21576: gap of unknown length

21577 41450: contig of 19874 bp in length
41451 41550: gap of unknown length
41551 55298: contig of 13748 bp in length
55299 55398: gap of unknown length
55399 66764: contig of 11366 bp in length
66765 66864: gap of unknown length
66865 81567: contig of 14703 bp in length
81568 81667: gap of unknown length
81668 94373: contig of 12706 bp in length
94374 94473: gap of unknown length
94474 104139: contig of 9666 bp in length
104140 104239: gap of unknown length
104240 114453: contig of 10214 bp in length
114454 114553: gap of unknown length
114554 122925: contig of 8372 bp in length
122926 123025: gap of unknown length
123026 127231: contig of 4206 bp in length
127232 127331: gap of unknown length
127332 134754: contig of 7423 bp in length
134755 134854: gap of unknown length
134855 140454: contig of 5600 bp in length
140455 140535: gap of unknown length
140536 147535: contig of 6981 bp in length
147536 147635: gap of unknown length
147636 153069: contig of 5434 bp in length
153070 153169: gap of unknown length
153170 159234: contig of 6065 bp in length
159235 159334: gap of unknown length
159335 164827: contig of 5493 bp in length
164828 164927: gap of unknown length
164928 169721: contig of 4794 bp in length
169722 169821: gap of unknown length
169822 172565: contig of 2744 bp in length
172566 172665: gap of unknown length
172666 175115: contig of 2450 bp in length
175116 175215: gap of unknown length
175216 177558: contig of 2543 bp in length
177559 177858: gap of unknown length
17859 180950: contig of 3092 bp in length
180951 181050: gap of unknown length
181051 184037: contig of 2987 bp in length
184038 184137: gap of unknown length
184138 187495: contig of 3358 bp in length
187496 187595: gap of unknown length
187596 189691: contig of 2096 bp in length
189692 189791: gap of unknown length
189792 191812: contig of 2021 bp in length
191813 191912: gap of unknown length
191913 194143: contig of 2231 bp in length.
location/Qualifiers
1. 194143
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-52513"

BASE COUNT 54237 a 42567 c 41198 g 53596 t 2545 others
ORIGIN
1. 194143
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-52513"

Query Match 44.3%; Score 435.8; DB 2; Length 194143;
Best Local Similarity 97.4%; Pred. No. 1.4e-113;
Matches 443; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

1 atgtgtgacccagtggtgtgacactggaagaatccccctgatgatgatacccaaac 60
|||||
Db 177208 ATGTGTGACCACTGAGTGAGTGAAGAAATCCCTGATGATGATGATCACTTAAC 177149
acatagaattcaaaatccctgctactacacacacacacatcatgtcttcaagagacccct 120
|||||
Db 177148 ACATGAAATTTCAAAATCCTGCTACTACACACACATCATGCTCTTCAAGACACCCCT 177089
gtctctcttcacaaagctctgtgaaagaagtggtcccaaggggctgggaat 180
|||||


```

Db 186895 ANGTGTAGCACACAGTGGGTGTGACTGTGAGAAATCCCTCATGATGATGACCTTAAC 186954
Oy 61 accatagaattcaaaatccctgcgtactacacacgaacatcgtcttaagaagaccct 120
Db 186955 ACCATAGAAATCAAAATCCCTGCTACTACACACGACATGTCCTTAAGAGCACCTCT 187014
Oy 121 gctctcttcacaaagctgtctgaacaaagattgtctccagaagggcctggggaat 180
Db 187015 GCTCTCTTCACACCAAGGCTGCTGAGAACAAAGATTGTCTCCAGAGGGGCTGGGGAT 187074
Oy 181 tttcagaacatgagtcacagagaggtgtcatgtgcctctgcagaatcccaatcag 240
Db 187075 TCTTCAGCAAAATGAGTCATGACAGAGGTGTCTCATGGCTTCAGAAATTCCTCAAT 187134
Oy 241 gagaagggcataacctgtgcagaagaaagtctctcttgaaagacattcttggaatag 300
Db 187135 GAGAAAGCCATTAACCTGTGCAAGAAAGATCTTCTTGGAAGCATCTCTTGAGAGAG 187194

RESULT 13
LOCUS HSA329005 699 bp DNA linear PRI 01-OCT-2001
DEFINITION Homo sapiens genomic sequence surrounding NotI site, clone
ACCESSION AJ329005
VERSION AJ329005.1 GI:15873423
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 699)
Kutsenko, A.S., Gizatullin, R.Z., Al-Amin, A.N., Wang, F.,
Podewski, R.M., Matsumkin, Y.G., Kvasha, S.M., Gyanchandani, A.,
Murevenko, O.V., Protodopov, A.I., Kashuba, V.I., Kisselev, L.L.,
Wasserman, M., Nahlestedt, C. and Zabarovsky, E.R.
Analysis of NotI flanking sequences: a new tool for gene discovery
and verification of the human genome
JOURNAL
REFERENCE
2 (bases 1 to 699)
Zabarovsky, E.R.
Direct Submission
Submitted (16-May-2001) Microbiology and Tumorbiology Centre,
Karolinska Institute, Theorells vag, 3, Box 280, Stockholm 171 77,
Sweden
FEATURES
source 1. 699
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NL1-FM8C"
BASE COUNT 178 a 164 c 171 g 185 t 1 others
ORIGIN
Query Match 25.8%; Score 253.8; DB 9; Length 699;
Best Local Similarity 99.2%; Pred. No. 1.4e-61;
Matches 255; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 673 agaaagctgaagaagaagcttgatgagccactccagatgagctgctcactct 732
Db 416 ACACAGCTGAAGAAAGATGAGCTTTGATGGGCCACTTCCAGGATGGCTGCTACTCT 357

```

```

Oy 733 gtctcaagaccatcacagaccaggtccctaattggtgtgagaccacaggggagaatcag 792
Db 356 GTTTTCAGAGACCATCACAGACAGGTCTTAATGGGTGTGTGACCCACAGGAGATCAGAG 297
Oy 793 gtcaagctcacagggtcttaaggtcctctgtataagatcagcagggcgaagctcacagct 852
Db 296 GTCAAGCTCAGGGGCTTTAAGGCTGCCCTTTAATAGACGTGACGGCCAAAGCTCACAGCT 237
Oy 853 atgcaacaccccgatgaacaggggtcctggtcctttgacacgaagtaactgaagaagac 912
Db 236 ATTACAAACCCACCGATGACAGAGGCTCTTGGCACCAGTAACCTGAAGAGAAC 177
Oy 913 tttcgcacatgatcca 929
Db 176 TTCTGCCATGATGATCA 160

RESULT 14
LOCUS AK000127 1731 bp mRNA linear PRI 22-FEB-2000
DEFINITION Homo sapiens cDNA FLJ20120 fis, clone COL05912.
ACCESSION AK000127
VERSION AK000127.1 GI:7020010
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens colon cDNA to mRNA, clone_lib:COL clone:COL05912.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (sites)
Kawabata, A., Hiki, J., Kobatake, N., Inagaki, H., Ikema, Y.,
Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishii, T.,
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
NEDO human cDNA sequencing project
JOURNAL
REFERENCE
2 (bases 1 to 1731)
Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishii, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
Direct Submission
Submitted (15-FEB-2000) to the DBJ/EMBL/GenBank databases. Sumio
Sugano, Institute of Medical Science, University of Tokyo, Department
of Virology, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: cdna@ems.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan: cDNA full insert
sequencing: Research Association for Biotechnology: cDNA library
construction: 5' - & 3' -end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source 1. 1731
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL05912"
/clone_lib="COL"
/tissue="colon"
/note="Cloning vector pMR18SFL3"
BASE COUNT 506 a 404 c 342 g 479 t
ORIGIN
Query Match 21.9%; Score 215.6; DB 9; Length 1731;
Best Local Similarity 98.2%; Pred. No. 1.3e-50;
Matches 218; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Oy 733 gtctcaagaccatcacagaccaggtccctaattggtgtgagaccacaggggagaatcag 792
Db 1 GTTTTCAGAGACCATCACAGACAGGTCTTAATGGGTGTGTGACCCACAGGAGATCAGAG 60
Oy 793 gtcaagctcacagggtcttaaggtcctctgtataagatcagcagggcgaagctcacagct 852
Db 1 GTTTCAGAGACCATCACAGACAGGTCTTAATGGGTGTGTGACCCACAGGAGATCAGAG 60

```

Db 61 GTCGAAGCTAGGCGCTTTAAGGCTGCGCTTGTATAGACGTACGCGCCACAGCTACAGCT 120

Qy 853 attgacacaccaccgcatgacacaggtctcctggccttggcaccacagctacctgaagaagac 912

Db 121 ATTGACACACACCGCATGACAGAGGCTCCTGGCGCTTTGGCACCACAGTACTGAAGAGAAC 180

Qy 913 ttctgcacatggatccgacgacgacgctggtgaggaataa 954

Db 181 TTCTCGCCATGATTCACAGACGAGGTGATGGTAAAGCGTA 222

RESULT 15

AC097248/LOCUS AC097248 149350 bp DNA linear HTG 20-DEC-2001

DEFINITION Rattus norvegicus clone CH230-177C11, *** SEQUENCING IN PROGRESS

AC097248

AC097248.3 GI:17973829

WORDS HTG: HTGS_PHASE1.

ORGANISM Rattus norvegicus

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

1 (bases 1 to 149350)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alspbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Biewe,K., Blankenburg,K., Bonni,D., Bouck,J., Bowie,S., Brivaga,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Demn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Hollway,C., Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratoch,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtharge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenwo,S., Oguh,M., Okunolu,G., Oragunye,N., Owiedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Prins,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojudoakan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shoochbat,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinsan,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,Y., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

TITLE Unpublished

JOURNAL 2 (bases 1 to 149350)

REFERENCE Worley,K.C.

AUTHORS Direct Submission

JOURNAL Submitted (13-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Dec 20, 2001 this sequence version replaced gi:17064657.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GINL

Center clone name: CH230-177C11

----- Summary Statistics

Assembly program: Phrap; version 0.990329First call to findPhrapList

Consensus quality: 114307 bases at least Q40

Consensus quality: 121845 bases at least Q30

Consensus quality: 128763 bases at least Q20

Estimated insert size: 114370; sum-of-contrigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 1.5x in Q20 bases; sum-of-contrigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 64 contrigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contrigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1

5540 5539: contrig of 5539 bp in length

5539: gap of unknown length

10594: contrig of 4953 bp in length

10594: gap of unknown length

10695: contrig of 5871 bp in length

10695: gap of unknown length

16665: contrig of 5234 bp in length

16665: gap of unknown length

21899: contrig of 4848 bp in length

21899: gap of unknown length

22000 26847: contrig of 4848 bp in length

26847: gap of unknown length

26948 30995: contrig of 4048 bp in length

30995: gap of unknown length

30996 31096: contrig of 4810 bp in length

31096: gap of unknown length

35906 36005: gap of unknown length

36005: contrig of 4016 bp in length

40022 40121: gap of unknown length

40121: contrig of 5067 bp in length

45189 45288: gap of unknown length

45288: contrig of 4171 bp in length

45289 49559: gap of unknown length

49559: contrig of 2237 bp in length

49560 51796: gap of unknown length

51796: contrig of 2966 bp in length

51897 54862: gap of unknown length

54862: contrig of 3191 bp in length

54863 54953: gap of unknown length

54953: contrig of 2490 bp in length

58154 58253: gap of unknown length

58253: contrig of 2832 bp in length

60744 60843: gap of unknown length

60843: contrig of 2624 bp in length

63775 63775: gap of unknown length

63775: contrig of 3368 bp in length

63776 66499: gap of unknown length

66499: contrig of 1760 bp in length

66500 69868 72345: gap of unknown length

72345: contrig of 1776 bp in length

72346 73890: gap of unknown length

73890: contrig of 1445 bp in length

73891 73990: gap of unknown length

73990: contrig of 2140 bp in length

76131 76230: gap of unknown length

76230: contrig of 1760 bp in length

77991 78090: gap of unknown length

78090: contrig of 1776 bp in length

78091 79966: gap of unknown length

79966: contrig of 2122 bp in length

79967 82088: contrig of 2122 bp in length

```

* 82089 82188: gap of unknown length
* 82189 84156: contig of 1968 bp in length
* 84157 84256: gap of unknown length
* 84257 86357: contig of 2101 bp in length
* 86358 86457: gap of unknown length
* 86458 87963: contig of 1506 bp in length
* 87964 88063: gap of unknown length
* 88064 89282: contig of 1219 bp in length
* 89283 91047: contig of 1665 bp in length
* 91048 91147: gap of unknown length
* 91148 93054: contig of 1907 bp in length
* 93055 93154: gap of unknown length
* 93155 95177: contig of 2023 bp in length
* 95178 95277: gap of unknown length
* 95278 96910: contig of 1633 bp in length
* 96911 97010: gap of unknown length
* 97011 98848: contig of 1838 bp in length
* 98849 98948: gap of unknown length
* 98949 100455: contig of 1507 bp in length
* 100456 102754: contig of 2199 bp in length
* 102755 102854: gap of unknown length
* 102855 104635: contig of 1781 bp in length
* 104636 104736: gap of unknown length
* 104736 106421: contig of 1685 bp in length
* 106421 106520: gap of unknown length
* 106521 107623: contig of 1242 bp in length
* 107623 107862: gap of unknown length
* 107863 109530: contig of 1668 bp in length
* 109531 109630: gap of unknown length
* 109631 111109: contig of 1479 bp in length
* 111110 111210: gap of unknown length
* 111210 113098: contig of 1889 bp in length
* 113099 113198: gap of unknown length
* 113199 114273: contig of 1075 bp in length
* 114274 114373: gap of unknown length
* 114373 116703: contig of 2330 bp in length
* 116704 116803: gap of unknown length
* 116804 117940: contig of 1137 bp in length
* 117941 118040: gap of unknown length
* 118041 120245: contig of 2205 bp in length
* 120246 120345: gap of unknown length
* 120346 121494: contig of 1149 bp in length
* 121495 121594: gap of unknown length
* 121595 123183: contig of 1589 bp in length
* 123184 123283: gap of unknown length
* 123284 124580: contig of 1297 bp in length
* 124581 124680: gap of unknown length
* 124681 126245: contig of 1565 bp in length
* 126246 126345: gap of unknown length
* 126346 127733: contig of 1388 bp in length
* 127734 127833: gap of unknown length
* 127834 129130: contig of 1297 bp in length
* 129131 129230: gap of unknown length
* 129231 130800: contig of 1570 bp in length
* 130801 130900: gap of unknown length
* 130901 132501: contig of 1601 bp in length
* 132502 132601: gap of unknown length
* 132602 134552: contig of 1951 bp in length
* 134553 134652: gap of unknown length
* 134653 135941: contig of 1289 bp in length
* 135942 136041: gap of unknown length
* 136042 137656: contig of 1615 bp in length
* 137657 137756: gap of unknown length
* 137757 138840: contig of 1084 bp in length
* 138841 138940: gap of unknown length

```

Query Match 17.9%; Score 176; DB 2; Length 149350;
 Best Local Similarity 71.5%; Pred. No. 4, 2e-39;
 Matches 289; Conservative 0; Mismatches 105; Indels 10; Gaps 4;

```

Db 96479 CCCCCGAGGCGATGATGACGACACGATAGAGTTCAAAAATCCGCTTTATGCCA 96420
QY 95 gacatcgtctcagaagacacccgtctctctccaccaagctgtgagacaagaa 154
Db 96419 GACACCATGCTTCAAGAGACCCCGGTGTCTTCCCAAGCTCTCCAGGACCGAA 96360
QY 155 gttgtccagaaggcgctgggaattgttcagaatgatagtacagaggtgcat 214
Db 96359 GTTGTCCAGAAACACACTGGGGACTTGT--CAACTGACTCTGACACAGATATCGT 96303
QY 215 ggccttcagaaattcccaatccagtlgagaaggcataaacttggcaagaaagtctt 274
Db 96302 TGCCCTGTAGAGATTCCCTCTCCAGTGAAGACATTAAGCTTGGCCAAAGAGTCTT 96243
QY 275 ctggaaagcattcttggagt--agtggagaaggaagtctgcagaagcgtcgca 331
Db 96242 CTTGAGAGACACTCTTCAGAGTAAACAGAGAGGAGGAGCCGCCAGCTCCCAAAAGG 96183
QY 332 agtctctgctcagaaggtcaaaagacgttggaatlacaaagatcgacagcagagtgt 391
Db 96182 AGATCCATGCTCAGGGGCCCTTCCCGGTAGAGGGGAGGTAGG--AACCAAGACTGTGT 96126
QY 392 ccaggtgtcttcttaacgttgagcagtgcttgagcatgaaact 435
Db 96125 CCAGATCCCTGTCCAGCGGTGAGAGCACTGCTGAGAGAGTGAAGCT 96082

```

Search completed: June 20, 2002, 05:05:07
 Job time: 9107 sec

QY 36 cccctagatgatgatacctaaccacatagaatc-aaaatcctcgctactacacca 94

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2002, 03:25:56 ; Search time 302.61 Seconds
(without alignments)
5582.908 Million cell updates/sec

Title: US-09-771-961-1
Perfect score: 984
Sequence: 1 atgttgacaccagtcggtggtg.....cacatgaagaagtactga 984

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	984	100.0	984	AAAD13235	Human BCL-X-like p
2	982.4	99.8	1179	AAH22582	Human BCL-G1 polyp
3	821	83.4	2132	AAAD13237	Human BCL-X-like p
4	791.6	80.4	1917	AAAS90453	DNA encoding novel
5	678	68.9	759	AAAD13236	Human BCL-X-like p
6	676.4	68.7	954	AAH22583	Human BCL-Gs polyp
7	505.2	51.3	987	AAH22670	Mouse Bcl-G polype
8	429	43.6	632	AAAL01348	Human reproductive
9	388	39.4	388	ABAA7100	Human breast cell

10	388	39.4	388	22	ABAA64982	Human foetal liver
11	388	39.4	388	22	ABAA32090	Probe #10556 for g
12	388	39.4	388	22	AAK13408	Human brain expres
13	388	39.4	388	22	AAK39145	Human bone marrow
14	388	39.4	388	22	AAI19954	Probe #9887 for ge
15	388	39.4	388	22	AAI45152	Probe #13838 used
16	388	39.4	388	22	AAI05666	Probe #5657 used t
17	317	32.2	466	22	AAH22582	Human foetal liver
18	317	32.2	466	22	ABAA2153	Probe #619 for gen
19	317	32.2	466	22	AAK00629	Human brain expres
20	317	32.2	466	22	AAK26079	Human bone marrow
21	317	32.2	466	22	AAI10708	Probe #641 for gen
22	317	32.2	466	22	AAI13196	Probe #652 used to
23	317	32.2	466	22	AAI00638	Probe #629 used to
24	272.4	27.7	8922	22	AAK87069	Human immune/haema
25	39.8	4.0	43	22	AAH22593	Bcl-Gs mutagenic p
26	39.8	4.0	43	22	AAH22594	Bcl-Gs mutagenic p
27	37.6	3.8	7055	20	AAK26303	Sequence of phage
28	37.6	3.8	7783	20	AAK26302	Sequence of phage
29	36.4	3.7	630	22	AAK31484	Porcine reproducti
30	36.2	3.7	759	22	AAAD13236	Human BCL-X-like p
31	36.2	3.7	954	22	AAH22583	Human BCL-Gs polyp
32	36.2	3.7	984	22	AAAD13235	Human BCL-X-like p
33	36.2	3.7	1179	22	AAH22582	Human BCL-G1 polyp
34	36.2	3.7	2132	22	AAAD13237	Human BCL-X-like p
35	35.8	3.6	10377	24	ABL34420	Human immune syste
36	35.6	3.6	484	20	AAV87011	EST clone AY93. H
37	35.2	3.6	6971	20	AAK26304	Sequence of phage
38	34.8	3.5	600	22	ABAA60917	Human foetal liver
39	34.8	3.5	600	22	ABAA28894	Probe #7360 for ge
40	34.8	3.5	600	22	AAK09207	Human brain expres
41	34.8	3.5	600	22	AAK35096	Human bone marrow
42	34.8	3.5	600	22	AAI40812	Probe #9498 used t
43	34.8	3.5	4528	17	AAI30055	S. pneumoniae dete
44	34.8	3.5	10711	19	AAV52278	Streptococcus pneu
45	34.6	3.5	1577	23	ABL27509	Drosophila melanog

ALIGNMENTS

RESULT 1	
AAAD13235	standard; cDNA: 984 BP.
XX	AAAD13235;
XX	23-OCT-2001 (first entry)
DT	Human BCL-X-like protein encoding cDNA #1.
DE	Human; BCL-X-like protein; therapy; physiological disorder; ss.
XX	Human;
XX	Homo sapiens.
OS	
XX	
FN	Key
FT	CDS
FT	Location/Qualifiers
FT	1..984
FT	/tag= a
FT	/product= "Human BCL-X-like protein #1"
XX	
PN	WO200157213-A2.
XX	
PD	09-AUG-2001.
XX	
PE	31-JAN-2001; 2001WO-US03446.
XX	
PR	04-FEB-2000; 2000US-0180412.
XX	
PA	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Donoho G, Hibun E, Turner CA, Friedrich G, Abuin A, Zambrowicz B,
XX	Sands AT;
XX	

DR WPI; 2001-488882/53.
DR P-PSDB; AAE07040.

XX Novel isolated human BCL-X-like polynucleotide, useful in therapeutic,
PT diagnostic and pharmacogenic applications
PS Claim 2a; Page 30; 33pp; English.

XX The present sequence is a cDNA encoding human BCL-X-like protein.
CC The BCL-X-like polynucleotides are useful in therapeutic, diagnostic
CC and pharmacogenic applications. They are useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic manifestations
CC perturbing the normal function of protein in the body and also for
CC treating physiological disorders and diseases. The BCL-X-like
CC polynucleotides are useful in conjunction with polymerase chain
CC reaction to screen libraries, isolate clones, to prepare cloning
CC and sequencing templates and as hybridisation probes for assessing
CC gene expression patterns.

XX Sequence 984 BP; 289 A; 230 C; 250 G; 215 T; 0 other;

XX Query Match 100.0%; Score 984; DB 22; Length 984;

XX Best Local Similarity 100.0%; Pred. NO. 6.6e-289;

XX Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgtagccacagtggtgtgacctggaagaataccccctagatgatgacctaac 60
DB 1 atgtgtagccacagtggtgtgacctggaagaataccccctagatgatgacctaac 60
QY 61 accatagaattcaaaatcctgcctactacacagacatcatgcttccaagacacct 120
DB 61 accatagaattcaaaatcctgcctactacacagacatcatgcttccaagacacct 120
QY 121 gctctcttcacacaaagcgtgtgagaagaagattgttcccaagggcctgggaat 180
DB 121 gctctcttcacacaaagcgtgtgagaagaagattgttcccaagggcctgggaat 180
QY 181 tgttcagcaaatgagtcagtcagagagtgctcatgacctgcagaataatcccaatcagt 240
DB 181 tgttcagcaaatgagtcagtcagagagtgctcatgacctgcagaataatcccaatcagt 240
QY 241 gagaagagccataaactctgtgcaagaagaagctctcttggaagaactctcttgagtagtg 300
DB 241 gagaagagccataaactctgtgcaagaagaagctctcttggaagaactctcttgagtagtg 300
QY 301 gagaagagccataaactctgtgcaagaagaagctctcttggaagaactctcttgagtagtg 360
DB 301 gagaagagccataaactctgtgcaagaagaagctctcttggaagaactctcttgagtagtg 360
QY 361 gaataccaagatctgcacagccagcagtggtcaggtgtcttcttaacgtgagcagtg 420
DB 361 gaataccaagatctgcacagccagcagtggtcaggtgtcttcttaacgtgagcagtg 420
QY 421 ttggagcatgaagctgtgagcccaaggtcatttccattgcacaacggagtagctgaatt 480
DB 421 ttggagcatgaagctgtgagcccaaggtcatttccattgcacaacggagtagctgaatt 480
QY 481 gtttattcctgcgacacacacaaagcagccagggaggttcaaggtccaaagagatt 540
DB 481 gtttattcctgcgacacacacaaagcagccagggaggttcaaggtccaaagagatt 540
QY 541 ttgttaactgaggtgtctctcttcacagctccaagcgccagtcgtgcttcaagttct 600
DB 541 ttgttaactgaggtgtctctcttcacagctccaagcgccagtcgtgcttcaagttct 600
QY 601 aagaagaatgaagaagaacaatactagccaataatgtgtgagctgctgaataattcaga 660
DB 601 aagaagaatgaagaagaacaatactagccaataatgtgtgagctgctgaataattcaga 660
QY 661 gatcagcttggaagaagaagctgaagaagaatgaagcttgatgggacattccaagatgg 720
DB 661 gatcagcttggaagaagaagctgaagaagaatgaagcttgatgggacattccaagatgg 720

QY 721 ctgtctactctgttttcaagaccatcacagaccaggtctctaattggtgtgaccacca 780
DB 721 ctgtctactctgttttcaagaccatcacagaccaggtctctaattggtgtgaccacca 780
QY 781 ggaagaatcagaagtcacaagctcaaggctttaaggctgccttgaatagacgacggcc 840
DB 781 ggaagaatcagaagtcacaagctcaaggctttaaggctgccttgaatagacgacggcc 840
QY 841 aagccacagcatttgacacacacccagatgaagaaggctccttggttgacacaaatgac 900
DB 841 aagccacagcatttgacacacacccagatgaagaaggctccttggttgacacaaatgac 900
QY 901 ctgaagaagaactctgcgcattgacccagcagcagcagtgatgagggaaaaatactggg 960
DB 901 ctgaagaagaactctgcgcattgacccagcagcagcagtgatgagggaaaaatactggg 960
QY 961 atatcatgaagaagtagactga 984
DB 961 atatcatgaagaagtagactga 984

RESULT 2

AAH22582
ID AAH22582 standard; cDNA; 1179 BP.

AAH22582;

07-SEP-2001 (first entry)

Human Bcl-G1 polypeptide encoding cDNA.

Bcl-G; cancer; cancer therapy; oncogene; apoptosis; Bcl-G1; cytosolic;

antiapoptotic; chromosome 12p12.3; human; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 196..1179

/tag= a

/product= "Bcl-G1"

13-DEC-2000; 2000WO-US33793.

14-DEC-1999; 99US-0461641.

(BURN-) BURNHAM INST.

Reed JC, Godzik A;

WPI; 2001-398125/42.

P-PSDB; AAB85166.

Novel polynucleotide encoding Bcl-G polypeptide, useful for modulating

apoptosis, and for diagnosing and treating cancer

Claim 3; Fig 1; 117pp; English.

The invention relates to Bcl-G polypeptides and nucleic acids encoding
CC them. The Bcl-G polypeptides can be expressed by standard recombinant
CC methodology. Bcl-G oligonucleotides (or its anti-sense strand) and Bcl-G
CC specific antibodies are useful for diagnosing cancer, monitoring cancer
CC therapy or assessing prognosis of patients with cancer. The Bcl-G
CC polypeptides are useful for modulating the activity of an oncogenic
CC polypeptide. They are useful for identifying modulators, for modulating
CC a level of apoptosis mediated by the Bcl-G polypeptide. A therapeutic
CC composition comprising the Bcl-G polypeptide, polynucleotide or antibody
CC is useful for treating a pathology characterized by abnormal cell
CC proliferation especially cancer. The present sequence represents a cDNA


```

OY 241 gagaaagcctaataaccttggcagaagaaagtcttcttggaaagcatcttggagtagtg 300
    |||||||
Db 432 gagaaagcctaataaccttggcagaagaaagtcttcttggaaagcatcttggagtagtg 491
OY 301 gagaaagcctaataaccttggcagaagaaagtcttcttggaaagcatcttggagtagtg 360
    |||||||
Db 492 gagaaagcctaataaccttggcagaagaaagtcttcttggaaagcatcttggagtagtg 551
OY 361 gaatacaagatcgcacacacacacacacacacacacacacacacacacacacacacac 420
    |||||||
Db 552 gaatacaagatcgcacacacacacacacacacacacacacacacacacacacacacac 611
OY 421 ttggagcatgaagctgttgagaccccaagtcattccattgcacaacggaatgaatt 480
    |||||||
Db 612 ttggagcatgaagctgttgagaccccaagtcattccattgcacaacggaatgaatt 671
OY 481 gttattctctggccacacacacacacacacacacacacacacacacacacacacacac 540
    |||||||
Db 672 gttattctctggccacacacacacacacacacacacacacacacacacacacacacac 731
OY 541 ttgttaactgaaggtctctctctccacacacacacacacacacacacacacacacacac 600
    |||||||
Db 732 ttgttaactgaaggtctctctctccacacacacacacacacacacacacacacacacac 791
OY 601 aagaaagatgaagaagaacaataactagccaataatgtgtgagctgtgtaaatatcagga 660
    |||||||
Db 792 aagaaagatgaagaagaacaataactagccaataatgtgtgagctgtgtaaatatcagga 851
OY 661 gatcagcttgagaaagaa----- 676
    |||||||
Db 852 gatcagcttgagaaagaaagacactgcctcatcccatcccttggttgacacacagcaltc 911
OY 677 ----- 676
    |||||||
Db 912 caggggtttccacagagatggttggatgctgcattggagctaaagaatgaactctctc 971
OY 677 -----agctgaagaa 687
    |||||||
Db 972 tgcctcgtgagcacaagctactgtaactgtaacttcttggtaacacagctgaagaa 1031
OY 688 gataagcgttgaatggccctccacacacacacacacacacacacacacacacacacacac 747
    |||||||
Db 1032 gataagcgttgaatggccctccacacacacacacacacacacacacacacacacacacac 1091
OY 748 aagaacacagctcctaattggtgtgagacccacagggagagaatcagaaggtcaaaagctcag 807
    |||||||
Db 1092 aagaacacagctcctaattggtgtgagacccacagggagagaatcagaaggtcaaaagctcag 1151
OY 808 tttaagcctgcctctgtaataagacgtcagcgcaagctcacagctatgacaacacccg 867
    |||||||
Db 1152 tttaagcctgcctctgtaataagacgtcagcgcaagctcacagctatgacaacacccg 1211
OY 868 atgaacacagctcctcgtggtcttgacacacacacacacacacacacacacacacacacac 927
    |||||||
Db 1212 atgaacacagctcctcgtggtcttgacacacacacacacacacacacacacacacacacac 1271
OY 928 cagcagcagcagctggtgagaaataatcttggagatcacacatgaagaagtgaactga 984
    |||||||
Db 1272 cagcagcagcagctggtgagaaataatcttggagatcacacatgaagaagtgaactga 1328

```

```

XX Homo sapiens.
OS
XX
PN W0200175067-A2.
XX
PD 11-OCN-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR P-PSDB; ABG26266.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID No 26257; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensic, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1917 BP; 553 A; 443 C; 439 G; 482 T; 0 other;

```

```

Query Match 80.4%; Score 791.6; DB 23; Length 1917;
Best Local Similarity 87.7%; Pred. No. 3.3e-238;
Matches 975; Conservative 0; Mismatches 9; Indels 128; Gaps 5;

```

```

OY 1 atgtgtgacacagctggtg-tgacctggaagaatccccctagatgatgaactaa 59
    |||||||
Db 96 atgtgtgacacagctggtg-tgacctggaagaatccccctagatgatgaactaa 155
OY 60 caccatgaattcaaaatcctgcctactacacacacacacacacacacacacacacacacacac 117
    |||||||
Db 156 caccatgaattcaaaatcctgcctactacacacacacacacacacacacacacacacacacac 215
OY 118 cctgctct-ccttcacacacacacacacacacacacacacacacacacacacacacacac 176
    |||||||
Db 216 cctgctctccttcacacacacacacacacacacacacacacacacacacacacacacac 275
OY 177 gaattgtcagaacaatgagatgacacagaggtgcatgaccttgacagaatcccaatc 236
    |||||||
Db 276 gaattgtcagaacaatgagatgacacagaggtgcatgaccttgacagaatcccaatc 335
OY 237 cagtgagaagcctaataaccttggcagaagaaagtcttcttggaaagcatcttggagtag 296
    |||||||

```

RESULT 4
 AAS90453 standard; cDNA; 1917 BP.
 AAS90453; -
 13-FEB-2002 (first entry)
 DNA encoding novel human diagnostic protein #26257.
 Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

Db 336 cagtgagaagggccataaaccttgagcaagaaaaagctctcttcttgaaagcattctttagt 335
 QY 297 agtggagaagaagatttcgcagacacgcctgccaaggtctctgtccaggttcaagagac 336
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 396 agtggagaagaagatttcgcagacacgcctgccaaggtctctgtccaggttcaagagac 455
 QY 357 gttggaatacagaatttcgcagacacgcctgccaaggtctctgtccaggttcaagagac 416
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 456 gttggaatacagaatttcgcagacacgcctgccaaggtctctgtccaggttcaagagac 515
 QY 417 gttggaatacagaatttcgcagacacgcctgccaaggtctctgtccaggttcaagagac 432
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 516 gttggaatacagaatttcgcagacacgcctgccaaggtctctgtccaggttcaagagac 575
 QY 433 gttggaatacagaatttcgcagacacgcctgccaaggtctctgtccaggttcaagagac 432
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 576 aaacctgtccaaaagccacatcattcttcaaggtctctgtccaggttcaagagac 635
 QY 433 gttggaatacagaatttcgcagacacgcctgccaaggtctctgtccaggttcaagagac 473
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 636 tagtctcaaatgttacaatgtctgagaccccaagtcattcttccatgtccacccagtagc 695
 QY 474 tgaatattgttattccctgcccacacccacacacacacacacacacacacacacacacac 533
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 696 tgaatattgttattccctgcccacacccacacacacacacacacacacacacacacacac 755
 QY 534 aggaatttttgaattcagaggtctctcttccacacacacacacacacacacacacacacac 593
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 756 aggaatttttgaattcagaggtctctcttccacacacacacacacacacacacacacacac 815
 QY 594 aagttctaaagaagaatgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 652
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 816 aagttctaaagaagaatgaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 875
 QY 653 atccagaagaatcagttggaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 712
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 876 atccagaagaatcagttggaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 935
 QY 713 aggaatttttgaattcagaggtctctcttccacacacacacacacacacacacacacacac 772
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 936 aggaatttttgaattcagaggtctctcttccacacacacacacacacacacacacacacac 995
 QY 773 accccaagggagaatcagaaggtcagaaggtcagaaggtcagaaggtcagaaggtcagaag 832
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 996 accccaagggagaatcagaaggtcagaaggtcagaaggtcagaaggtcagaaggtcagaag 1055
 QY 833 tcaagcccaaggtcagaaggtcagaaggtcagaaggtcagaaggtcagaaggtcagaaggt 892
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1056 tcaagcccaaggtcagaaggtcagaaggtcagaaggtcagaaggtcagaaggtcagaaggt 1115
 QY 893 ccaagttaccctgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 952
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1116 ccaagttaccctgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1175
 QY 953 tacttgggatacacatgagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 984
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1176 tacttgggatacacatgagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 1207
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 RESULT 5
 AAD13236 standard; cDNA; 759 BP.
 XX AAD13236:
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Human BCL-X-like protein encoding cDNA #2.
 XX
 KW Human; BCL-X-like protein; therapy; physiological disorder; ss.
 XX
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT CDS 1..759
 FT /tag= a
 FT /product= "Human BCL-X-like protein #2"
 XX
 XX W0200157213-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 31-JAN-2001; 2001WO-US03446.
 XX
 XX 04-FEB-2000; 2000US-0180412.
 XX
 XX (LEXI-) LEXICON GENETICS INC.
 XX
 XX Donoho G, Hilbun E, Turner CA, Friedrich G, Abuin A, Zambrowicz B;
 PI Sands AT;
 XX
 XX WPI; 2001-488882/53.
 DR P-PSDB; AAE07041.
 XX
 PT Novel isolated human BCL-X-like polynucleotide, useful in therapeutic,
 PT diagnostic and pharmacogenic applications
 XX
 XX Disclosure; Page 31; 33pp; English.
 XX
 XX The present sequence is a cDNA encoding human BCL-X-like protein.
 CC The BCL-X-like polynucleotides are useful in therapeutic, diagnostic
 CC and pharmacogenic applications. They are useful for screening drugs
 CC effective in the treatment of symptomatic or phenotypic manifestations
 CC perturbing the normal function of protein in the body and also for
 CC treating physiological disorders and diseases. The BCL-X-like
 CC polynucleotides are useful in conjunction with polymerase chain
 CC reaction to screen libraries, isolate clones, to prepare cloning
 CC and sequencing templates and as hybridisation probes for assessing
 CC gene expression patterns.
 XX
 XX Sequence 759 BP; 216 A; 182 C; 185 G; 176 T; 0 other;
 XX
 Query Match 68.9%; Score 678; DB 22; Length 759;
 Best Local Similarity 100.0%; Pred. No. 1.1e-202;
 Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 atgtctagaccacagtggtgagccttggaagaatccccctagatgatgacctaacc 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 atgtctagaccacagtggtgagccttggaagaatccccctagatgatgacctaacc 60
 QY 61 accatagaattcaaaaatcctcgcctactacacacacacacacacacacacacacacacac 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 accatagaattcaaaaatcctcgcctactacacacacacacacacacacacacacacacac 120
 QY 121 gctctcttccaccaaagctgctgagagaagaagttgtccagaagggcctggggaat 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 gctctcttccaccaaagctgctgagagaagaagttgtccagaagggcctggggaat 180
 QY 181 ttttcaagcaaatgagtcagagacagaggtgtcatgagccttcagaaattcccaatccaggt 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 ttttcaagcaaatgagtcagagacagaggtgtcatgagccttcagaaattcccaatccaggt 240
 QY 241 gagaagggccataaaccttggccaagaaagtctctcttggaagaagattcttggagtagtg 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 gagaagggccataaaccttggccaagaaagtctctcttggaagaagattcttggagtagtg 300
 QY 301 gagaaggaagatttcgacagacacgcctgccaaggtctctctgcaaggtcacaagagcgttg 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 gagaaggaagatttcgacagacacgcctgccaaggtctctctgcaaggtcacaagagcgttg 360
 QY 361 gaataccaagatttcgacacacacgcctgccaaggtctctctgcaaggtcacaagagcgttg 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 gaataccaagatttcgacacacacgcctgccaaggtctctctgcaaggtcacaagagcgttg 420
 QY 421 ttggagcatgaagctgtggaccccaagtcattctcatctgccaacggagtagctgaatt 480

|||||
Db ttggagcatgaagctgtagcccaagatcattccattgccaacggtagctgaatt 480
QY 481 gtttattccggccaccacacacacagccagcagaggttcaatgccaaagatt 540
Db 481 gtttattccggccaccacacacacagcagcagaggttcaatgccaaagatt 540
QY 541 ttgttaactgaggggtctctctctcccaagccagctgagcttcaagttct 600
Db 541 ttgttaactgaggggtctctctctcccaagccagctgagcttcaagttct 600
QY 601 aagaaagatgaagaagaacaatactagccaaattgttgagctgtgaataatcaga 660
Db 601 aagaaagatgaagaagaacaatactagccaaattgttgagctgtgaataatcaga 660
QY 661 gatcagttggaagaagaag 678
Db 661 gatcagttggaagaagaag 678

AAH22583 standard; cDNA; 954 BP.
AAH22583

07-SEP-2001 (first entry)

Human Bcl-Gs polypeptide encoding cDNA.

Bcl-G; cancer; cancer therapy; oncogene; apoptosis; Bcl-Gs; cytosolic;
antiapoptotic; chromosome 12p12.3; human; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 196..954
FT /tag="a
FT /product="Bcl-Gs"

WO200144282-A2.

21-JUN-2001.

13-DEC-2000; 2000WO-US33793.

14-DEC-1999; 99US-0461641.

(BURN-) BURNHAM INST.

Reed JC, Godzik A;

WPI: 2001-398125/42.

P-PSDB: AAB85167.

Novel polynucleotide encoding Bcl-G polypeptide, useful for modulating
apoptosis, and for diagnosing and treating cancer

Claim 3; Fig 3; 117pp. English.

The invention relates to Bcl-G polypeptides and nucleic acids encoding
them. The Bcl-G polypeptides can be expressed by standard recombinant
methodology. Bcl-G oligonucleotides (or its anti-sense strand) and Bcl-G
specific antibodies are useful for diagnosing cancer, monitoring cancer
therapy or assessing prognosis of patients with cancer. The Bcl-G
polypeptides are useful for modulating the activity of an oncogenic
polypeptide. They are useful for identifying modulators, for modulating
a level of apoptosis mediated by the Bcl-G polypeptide. A therapeutic
composition comprising the Bcl-G polypeptide, polynucleotide or antibody
is useful for treating a pathology characterized by abnormal cell
proliferation especially cancer. The present sequence represents a cDNA
encoding a Bcl-Gs polypeptide.

Sequence 954 BP; 267 A; 237 C; 240 G; 210 T; 0 other;

Query Match 68.7%; Score 676.4; DB 22; Length 954;

Best Local Similarity 99.9%; Pred. No. 3.9e-202;

Matches 677; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgtgtaacacagatgggtgtgactctggagaagaatcccttaatatgatacctaacc 60
Db 196 atgtgtaacacagatgggtgtgactctggagaagaatcccttaatatgatacctaacc 255
QY 61 accatagaattccaataatctctgactactacacacagacatcatgtcttcaagaacacct 120
Db 256 accatagaattccaataatctctgactactacacacagacatcatgtcttcaagaacacct 315
QY 121 gctctctctccaacagctctgagaaagaagttttccagaggggcttgggaat 180
Db 316 gctctctctccaacagctctgagaaagaagttttccagaggggcttgggaat 375
QY 181 ttgtcagcaaatgagtcagacagaggtgtatggccttcgcaagaattcccaatcagt 240
Db 376 ttgtcagcaaatgagtcagacagaggtgtatggccttcgcaagaattcccaatcagt 435
QY 241 gagaagccataaaactctgccaagaagaagttctcttggaaagcatctcttggagtagtg 300
Db 436 gagaagccataaaactctgccaagaagaagttctcttggaaagcatctcttggagtagtg 495
QY 301 gagaaggaagattcgcaagacgcctcgcaaggtctctgctcaaggtcaagaagcgttg 360
Db 496 gagaaggaagattcgcaagacgcctcgcaaggtctctgctcaaggtcaagaagcgttg 555
QY 361 gaataccaagattcgcaagcagctgctcaaggtctcttctaagctggagcagtg 420
Db 556 gaataccaagattcgcaagcagctgctcaaggtctcttctaagctggagcagtg 615
QY 421 ttggaagcatgaagctgtgagaccacaaagtcatctccattgccaacccagtagctgaatt 480
Db 616 ttggaagcatgaagctgtgagaccacaaagtcatctccattgccaacccagtagctgaatt 675
QY 481 gtttattccctgcccacacacacacacagcccgagcaggggtcttcaagttccaaagatt 540
Db 676 gtttattccctgcccacacacacacacagcccgagcaggggtcttcaagttccaaagatt 735
QY 541 ttgttaactgaggggtctctctctccagctcgaagccagctgctgtagcttcaagttct 600
Db 736 ttgttaactgaggggtctctctctccagctcgaagccagctgctgtagcttcaagttct 755
QY 601 aagaaagatgaagaagaacaatactagccaaattgttgagctgtgaataatcaga 660
Db 796 aagaaagatgaagaagaacaatactagccaaattgttgagctgtgaataatcaga 855
QY 661 gatcagttggaagaagaag 678
Db 856 gatcagttggaagaagaag 873

RESULT 7

AAH22670 standard; DNA; 987 BP.

AAH22670;

07-SEP-2001 (first entry)

Mouse Bcl-G polypeptide encoding cDNA.

Bcl-G; cancer; cancer therapy; oncogene; apoptosis; cytosolic;

antiapoptotic; chromosome 12p12.3; mouse; ss.

Mus sp.

Key Location/Qualifiers
CDS 1..987

```

FT      /*tag= a
FT      /product= "Bcl-G"
FT      /transl_except= "(pos: 319..321, aa: Xaa)"
XX
XX
XX      WO200144282-A2.
XX
XX      21-JUN-2001.
XX
XX      13-DEC-2000; 2000MO-US33793.
XX
XX      14-DEC-1999; 99US-0461641.
XX
XX      (BURN-) BURNHAM INST.
XX
XX      Reed JC, Godzik A;
XX
XX      WPI; 2001-398125/42.
XX      P-PSDB; AAB85188.
XX
XX      Novel polynucleotide encoding Bcl-G polypeptide, useful for modulating
XX      apoptosis, and for diagnosing and treating cancer -
XX
XX      Claim 3; Page 114-115; 117p; English.
XX
XX      The invention relates to Bcl-G polypeptides and nucleic acids encoding
XX      them. The Bcl-G polypeptides can be expressed by standard recombinant
XX      methodology. Bcl-G oligonucleotides (or its anti-sense strand) and Bcl-G
XX      specific antibodies are useful for diagnosing cancer, monitoring cancer
XX      therapy or assessing prognosis of patients with cancer. The Bcl-G
XX      polypeptides are useful for modulating the activity of an oncogenic
XX      polypeptide. They are useful for identifying modulators, for modulating
XX      a level of apoptosis mediated by the Bcl-G polypeptide. A therapeutic
XX      composition comprising the Bcl-G polypeptide, polynucleotide or antibody
XX      is useful for treating a pathology characterized by abnormal cell
XX      proliferation especially cancer. The present sequence represents a
XX      mouse Bcl-G cDNA.
XX
XX      Sequence 987 BP; 272 A; 246 C; 271 G; 197 T; 1 other;

```

```

Query Match      51.3%; Score 505.2; DB 22; Length 987;
Best Local Similarity 72.9%; Pred. No. 2.9e-148;
Matches 728; Conservative 0; Mismatches 244; Indels 27; Gaps 5;

QY      1 atgttgagccacgtgggtgtgtgacctggaagaatccccctgatatgatgacctaaac 60
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      1 atgtgcagcaccagttgtatgacctgtgaagaacatccccctgagatgatgtccaac 60
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      61 accatagaattcaaatcctcgcctactacacacagacatcatgtcttcaagagaccct 120
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      61 agcatagagtcaaatctcgtcctcttaagcagacaccatgtcttcaagaacacccg 120
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      121 gctctcttctaccaaagctctgagaacaagaatttccacagaggcgctcgaggaaat 180
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      121 gctgtctctcgcacagctctccagaacaagagctctgccagaagaccctcgaggact 180
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      181 ttttcagcaaatgtagtgcagacagagtgtcatgacctgctgagaattcccaatcagat 240
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      181 tgtt---caactgttctctggaacacagatlatcatgtccttgcaagagttcccccctcagc 237
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      241 gagaagccataaactctgccaagaanaaagctctcttgaaagcatctcttgagtagtg 300
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      238 gaaaagaacatcagcttgccaagaagaagatctctcttggaacacactcttcagggtggcc 297
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      301 gaga---aggaagattcgcaagcgcgcctgcgaaggtcttcgtctaaggtc----- 349
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      298 gagaaggaaggaagcctcgcngctcccccagaagagatcgcagatcgaaggtcctcagggc 357
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      350 ----aaaggaagcttggaaatccaacaaattccgacagcagcagcagtgctcttct 405
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      358 cccctcccgtagagcgcgacagagtgtctccacaacacacacacacggtcccggtctctgagc 417
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY      406 aacgtgagcagtgcttgtagcatgaagctgtgaccccaagaattcttcattgccaac 465
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB      418 agtggagcag-ccttgagagtgaaattgttgattcccaagtggtcttgattgccaac 476
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      466 cgagtgcgtgaattgtttattctcttgccacacacacacacagcagccgagcgaggtctc 525
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      477 agagtgcgtgaattgtttattctcttgccacacacacacacagcagccgagcgagga 536
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      526 aagtcacaagagatttcttaactgaggtctctcttcacagctcccaagcagcgaggtcct 585
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      537 agccagctcaaaagagagggctctcgagagatttctgattcaggttttgaagacgttggga 596
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      586 gtagcttcaagttctaaagaagaatgaagaacaaatactagccaaatgttgaagctg 645
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      597 ct-----ctaagataaagatgtgtgaagaccaaataataaagcaagattgttgagctg 648
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      646 ctgaatatctcagggagatcagtttggaagaagcctggaagaagtaaggtcttgatggc 705
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      649 ctgaatctcgggggtatcagtttggaagaagataaagaagaagcagagcttgtaggc 708
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      706 cacttccagagtggtgtctctactctgtttccaagccatcacagccaggtcctaagt 765
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      709 agcttccagagcgggtgtctctactctcaagcttcaagacatcacagaccttctctgag 768
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      766 ggtgtgaccccaaggaggaatcagaagtgtaaaagctcagggcttgaagctgccttgta 825
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      769 gacgtggaacacagagagatcatcagagtgtaaaagctcggggttcaagctgccttgca 828
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      826 atagagcttaagcagcagctacagctatgtgacacacacccagatgaagaaggtctcggc 885
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      829 atagagcctatcgccaagctcagcagctcagcagcagcagcagcagcagcagcagcag 888
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      886 ttggcaccacaaatctcgaagaagaacctctcgccatgataccagcagcagcagcagcag 945
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      889 ttccggaccacaaatctcgaagaagaacctctcccccgttggttcgcaagatcgcgagcag 948
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY      946 gaaaaaatacttgggatalcatcagtaagaagttagctga 984
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      949 gaaaaaatacttgggatalcatcagtaagaagttagctga 987
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 8
AAL01348
ID      AAL01348 standard; cDNA; 632 BP.
XX
XX      AAL01348;
XX
XX      21-NOV-2001 (first entry)
XX
XX      Human reproductive system related antigen cDNA SEQ ID NO: 1349.
XX
XX      Human: reproductive system related antigen; reproductive system disorder;
XX      cancer; gene therapy; ss.
XX
XX      Homo sapiens.
XX
XX      WO200155320-A2.
XX
XX      02-AUG-2001.
XX
XX      17-JAN-2001; 2001WO-US01339.
XX
XX      31-JAN-2000; 2000US-0179065.
XX      04-FEB-2000; 2000US-0180628.
XX      24-FEB-2000; 2000US-0184664.
XX      02-MAR-2000; 2000US-0186350.
XX      16-MAR-2000; 2000US-0189874.
XX      17-MAR-2000; 2000US-0190676.
XX      18-APR-2000; 2000US-0198123.
XX      19-MAY-2000; 2000US-0205515.
XX      07-JUN-2000; 2000US-0209467.
XX      28-JUN-2000; 2000US-0214886.
XX      30-JUN-2000; 2000US-0215135.
XX      07-JUL-2000; 2000US-0216647.

```

PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0236807.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239835.
PR 13-OCT-2000; 2000US-0239837.
PR 20-OCT-2000; 2000US-0240560.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI: 2001-465570/50.
XX P-PsDB; AAM95378.
XX
XX
PT Isolated nucleic acid molecule encoding a reproductive system antigen
PT is used in preventing, treating or ameliorating a medical condition -
XX
XX
PS Claim 1; SEQ ID NO 1349; 1297bp + Sequence Listing; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of human reproductive system related antigens. These can be used
CC in the prevention and treatment of reproductive system disorders,
CC including cancer. The present sequence is a coding sequence of the
XX invention.
XX
SQ Sequence 632 BP; 190 A; 143 C; 166 G; 128 T; 5 other;

XX 01-FEB-2002 (first entry)
 XX Human foetal liver single exon nucleic acid probe #13287.
 DE Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX Homo sapiens.
 XX MO200157277-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00669.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483447/52.

Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human fetal liver -

Claim 4; SEQ ID NO 13287; 639pp + sequence listing; English.

The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a single exon nucleic acid probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;

Query Match 39.4%; Score 388; DB 22; Length 388;

Best Local Similarity 100.0%; Pred. No. 1.3e-111; Mismatches 0; Indels 0; Gaps 0;

46 gatgatgacctaaacacacataatcaaatcctgcctactacacagacatcatgctc
 1 gatgatgacctaaacacacataatcaaatcctgcctactacacagacatcatgctc
 106 ttcgaagagacacccctctctctctcaccaaagctgtgagaaagaagttgtccag
 61 ttcgaagagacacccctctctctctcaccaaagctgtgagaaagaagttgtccag
 166 aggggctgtggagatgttgcgaatgagatgcagagaggtgtcattgctgtcaga
 121 aggggctgtggagatgttgcgaatgagatgcagagaggtgtcattgctgtcaga
 226 aattcccaatccagtgagagagcaccataaccttggcaagaanaatctcttgaaga
 181 aattcccaatccagtgagagagcaccataaccttggcaagaanaatctcttgaaga
 286 ttccttggagtagtgagagagagatcgcagagacgcctgtgccaaggtctctcag
 241 ttccttggagtagtgagagagagatcgcagagacgcctgtgccaaggtctctcag
 346 ggtcaagagacgttgcgaatcgaagatcgcagagcagcagctgtgtccaggtcttct

Db 301 ggtcaagagacgttgcgaatcgaagatcgcagagcagcagctgtgtcttct 360
 Oy 406 aacgtgagacagtgcttgcagacatgaag 433
 Db 361 aacgtgagacagtgcttgcagacatgaag 388

RESULT 11

ABA32090
 ID ABA32090 standard; DNA; 388 BP.

AC ABA32090;

DT 23-JAN-2002 (first entry)

DE Probe #10556 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease; ss.

XX Homo sapiens.

XX MO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488899/53.

Single exon nucleic acid probes for analyzing gene expression in human hearts -

Claim 4; SEQ ID NO 10556; 530pp; English.

The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart. The

present sequence is one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived

from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging,

monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

congenital heart disease.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pct_sequences.

Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;

Query Match 39.4%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. No. 1.3e-111; Mismatches 0; Indels 0; Gaps 0;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 46 gatgatgacctaaacacacataatcaaatcctgcctactacacagacatcatgctc
 Db 1 gatgatgacctaaacacacataatcaaatcctgcctactacacagacatcatgctc
 Oy 106 ttcgaagagacacccctctctctcaccaaagctgtgagaaagaagttgtccag 165

Db 61 ttcaagagcaccctgctctctctctcaccagaagctgctgagacaagaattgtccag 120
 QY 166 aggggctgggggagattgttaagcaaatgacatgacagaggttcaatgcttgcaga 225
 Db 121 aggggctgggggagattgttcaagcaaatgacatgacagaggttcaatgcttgcaga 180
 QY 226 aattcccaatcccaatgaggaagccataaacttgcagaagaagaatcttcttgcaga 285
 Db 181 aattcccaatcccaatgaggaagccataaacttgcagaagaagaatcttcttgcaga 240
 QY 286 ttcttggagtagtggagaaagaaatctgcagagcagcctgccaaggtcttgcag 345
 Db 241 ttcttggagtagtggagaaagaaatctgcagagcagcctgccaaggtcttgcag 300
 QY 346 ggtcaaaagacgttggaaataccaagattcgacagccagcagtggtccaggtcttct 405
 Db 301 ggtcaaaagacgttggaaataccaagattcgacagccagcagtggtccaggtcttct 360
 QY 406 aacgtgagcagtgcttggagcatgaag 433
 Db 361 aacgtgagcagtgcttggagcatgaag 388

RESULT 12
 ID AAK13408
 AAK13408 standard; DNA; 388 BP.

AC AAK13408;
 DT 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe SEQ ID NO: 13399.
 XX Human; brain expressed exon; gene expression analysis; probe;
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KM epilepsy; cancer; ss.
 OS Homo sapiens.

WO200157275-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00667.
 04-FEB-2000; 2000US-0180312.
 26-MAY-2000; 2000US-0207456.
 30-JUN-2000; 2000US-0608408.
 03-AUG-2000; 2000US-0632366.
 21-SEP-2000; 2000US-0234687.
 27-SEP-2000; 2000US-0236359.
 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;
 WPI; 2001-483446/52.

Single exon nucleic acid probes for analyzing gene expression in human brains -

Example 4; SEQ ID NO: 13399; 650pp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention.

SQ Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;

Query Match 39.4%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. No. 1.3e-111;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 gatgatgacctaaacaacatagaaatcaaatctctgctactacacagacatcatgtc 105
 Db 1 gatgatgacctaaacaacatagaaatcaaatctctgctactacacagacatcatgtc 60
 QY 106 ttcaagagcaccctgctctcttctcaccagaagctgctggaacaagaatttgcag 165
 Db 61 ttcaagagcaccctgctctcttctcaccagaagctgctggaacaagaatttgcag 120
 QY 166 aggggctgggggagattgttcaagcaaatgacatgacagaggttcaatgcttgcaga 225
 Db 121 aggggctgggggagattgttcaagcaaatgacatgacagaggttcaatgcttgcaga 180
 QY 226 aattcccaatcccaatgaggaagccataaacttgcagaagaagaatcttcttgcaga 285
 Db 181 aattcccaatcccaatgaggaagccataaacttgcagaagaagaatcttcttgcaga 240
 QY 286 ttcttggagtagtggagaaagaaatctgcagagcagcctgccaaggtcttgcag 345
 Db 241 ttcttggagtagtggagaaagaaatctgcagagcagcctgccaaggtcttgcag 300
 QY 346 ggtcaaaagacgttggaaataccaagattcgacagccagcagtggtccaggtcttct 405
 Db 301 ggtcaaaagacgttggaaataccaagattcgacagccagcagtggtccaggtcttct 360
 QY 406 aacgtgagcagtgcttggagcatgaag 433
 Db 361 aacgtgagcagtgcttggagcatgaag 388

RESULT 13
 ID AAK39145
 AAK39145 standard; DNA; 388 BP.

AC AAK39145;
 DT 06-NOV-2001 (first entry)
 DE Human bone marrow expressed single exon probe SEQ ID NO: 13702.

XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; leukemia; lymphoma; myeloma; ss.

OS Homo sapiens.

WO200157276-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US00668.

04-FEB-2000; 2000US-0180312.
 26-MAY-2000; 2000US-0207456.
 30-JUN-2000; 2000US-0608408.
 03-AUG-2000; 2000US-0632366.
 21-SEP-2000; 2000US-0234687.
 27-SEP-2000; 2000US-0236359.
 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;
 WPI; 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 13702; 658bp + Sequence Listing; English.
 PS The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 XX
 SO Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;

Query Match 39.4%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. No. 1.3e-111;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 gatgtgacctaaacccatagaattcaaatcctgcctactacacagacatcatgtc 105
 Db 1 gatgatgacctaaacccatagaattcaaatcctgcctactacacagacatcatgtc 60
 106 ttcaagagacacccctctctctctccacaaagctgtgagaaagaagttgtccag 165
 61 ttcaagagacacccctctctctctccacaaagctgtgagaaagaagttgtccag 120
 QY 166 aggggcttgagggaattgttcagcaaatgatgatgagacagagtgatgacctgcaga 225
 Db 121 aggggcttgagggaattgttcagcaaatgatgatgagacagagtgatgacctgcaga 180
 QY 226 aattcccaatcagtgagagaagggccataaactctgagcaagaagaagttcttggaaaga 285
 Db 181 aattcccaatcagtgagagaagggccataaactctgagcaagaagaagttcttggaaaga 240
 QY 286 ttctttgagtagtgagagaaggaagattcgacagacagcctgcacaaagtctctgtcag 345
 Db 241 ttctttgagtagtgagagaaggaagattcgacagacagcctgcacaaagtctctgtcag 300
 QY 346 ggtcaaaagagcgttggaatacaagaattcgacacagcagcagtggtccagtgcttct 405
 Db 301 ggtcaaaagagcgttggaatacaagaattcgacacagcagcagtggtccagtgcttct 360
 QY 406 aacgttgagcagtgcttgagcatgaag 433
 Db 361 aacgttgagcagtgcttgagcatgaag 388

RESULT 14
 AAI19954
 ID AAI19954 standard; DNA; 388 BP.
 XX AAI19954;
 DT 12-OCT-2001 (first entry)
 XX
 DE Probe #9887 for gene expression analysis in human cervical cell sample.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 XX
 OS Homo sapiens.
 XX
 XX MO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001; 2001MO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488901/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 PS
 PS Claim 25; SEQ ID No 9887; 487bp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENPs). The present sequence is one such probe. The SENPs are derived
 CC from human Hela cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging
 CC of diseases of the cervix, notably cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;

Query Match 39.4%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. No. 1.3e-111;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 gatgatgacctaaacacacatagaattcaaatcctgcctactacacagacatcatgtc 105
 Db 1 gatgatgacctaaacacacatagaattcaaatcctgcctactacacagacatcatgtc 60
 QY 106 ttcaagagacacccctctctctctccacaaagctgtgagaaagaagttgtccag 165
 Db 61 ttcaagagacacccctctctctctccacaaagctgtgagaaagaagttgtccag 120
 QY 166 aggggcttgagggaattgttcagcaaatgatgatgagacagagtgatgacctgcaga 225
 Db 121 aggggcttgagggaattgttcagcaaatgatgatgagacagagtgatgacctgcaga 180
 QY 226 aattcccaatcagtgagagaagggccataaactctgagcaagaagaagttcttggaaaga 285
 Db 181 aattcccaatcagtgagagaagggccataaactctgagcaagaagaagttcttggaaaga 240
 QY 286 ttctttgagtagtgagagaaggaagattcgacagacagcctgcacaaagtctctgtcag 345
 Db 241 ttctttgagtagtgagagaaggaagattcgacagacagcctgcacaaagtctctgtcag 300
 QY 346 ggtcaaaagagcgttggaatacaagaattcgacacagcagcagtggtccagtgcttct 405
 Db 301 ggtcaaaagagcgttggaatacaagaattcgacacagcagcagtggtccagtgcttct 360
 QY 406 aacgttgagcagtgcttgagcatgaag 433
 Db 361 aacgttgagcagtgcttgagcatgaag 388

RESULT 15
 AAI45152
 ID AAI45152 standard; DNA; 388 BP.
 XX AAI45152;
 AC
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #13838 used to measure gene expression in human placenta sample.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2002, 02:32:30 ; Search time 2402.36 Seconds
(without alignments)
5528.315 Million cell updates/sec

Title: US-09-771-961-1

Perfect score: 984

Sequence: 1 atgctagcaccagtggtg.....cacatgaagaagtagactga 984

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

al number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estbba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_iny:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	694.4	70.6	980	10	BI833102 603090814
2	603.4	61.3	763	10	BG178587 602696672
3	549	55.8	774	10	BG720221 602692271
4	527	53.6	1535	11	AK016997 Mus muscu
5	527	53.6	1541	11	AK008682 Mus muscu
6	527	53.6	1578	11	AK018579 Mus muscu
7	527	53.6	2232	11	AK016670 Mus muscu
8	506.6	51.5	715	10	BI838169 603083677
9	497.2	50.5	672	10	BG177835 602693855
10	418.4	42.5	703	10	BG178835 602693855
11	376	38.2	426	9	AM000827 wu45d10. x
12	359	36.5	676	10	BI652080 603298866
13	356.8	36.3	422	12	AO131601 HS 3049_A
14	342.6	34.8	984	9	BB013694 BB013694
15	321.4	32.7	406	9	AA399486 z660c07.r
16	303	30.8	451	10	BM433532 BM433532
17	276.2	28.1	776	10	BI463753 603203445

18	255.8	26.0	394	10	BE631990
19	246.6	25.1	446	9	BB746476
20	246.6	25.1	920	10	BI453888
21	246.4	25.0	652	9	AA536718
22	243.6	24.8	680	10	BG245730
23	242.6	24.7	448	10	BF198484
24	234.4	23.8	697	10	BF584699
25	227.2	23.1	924	10	BF176106
26	209	21.2	666	9	BB628327
27	204.4	20.8	644	10	BM211865
28	203.8	20.7	1061	10	BF139422
29	201.2	20.4	403	10	BI342342
30	197	20.0	567	9	AM631581
31	190.6	19.4	922	9	BB615307
32	186	18.9	465	10	BF149673
33	149	15.1	423	10	BG895661
34	141.8	14.4	559	10	BM216619
35	133.6	13.6	437	9	AA061006
36	132	13.4	848	10	BI158784
37	130.6	13.3	397	9	AI614194
38	123	12.5	354	9	AM414655
39	109.2	11.1	376	9	AM484860
40	104.8	10.7	370	9	AM428115
41	97.6	9.9	216	10	BI360692
42	93.8	9.5	332	10	T69681
43	88.6	9.0	473	9	AA400686
44	87.4	8.9	256	9	AI606692
45	83	8.4	467	9	AM001213

ALIGNMENTS

RESULT 1
BI833102
LOCUS
DEFINITION
603090814F1 NIH_MGC_120 Homo sapiens CDNA clone IMAGE:5229819 5',
mRNA sequence.
BI833102
ACCESSION
BI833102.1 GI:15944652
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM11578 row: k column: 04
High quality sequence stop: 803.

FEATURES

source
1..980
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5229819"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-Sport6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male; Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is

Db	565	TTGGAGCATGAAGCTGTGGACCCCAATTCATTTTCATTGGCCAAACCGAGTACTGAATT	624
Qy	481	gttattcctggccaccacacaaagcgaaccagcaggaggtcctaagtlccaaagagalt	540
Db	625	GTTTATCTCTGGCCACACACCAACGAGCCACGAGGAGGCTTCAGTCCAAAGAGATT	684
Qy	541	tttgaactgagggctctccttccttcagctccaagcgcgycctgtatgtcctaagtcct	600
Db	685	TTTGTAACTGAAGGCTCTCTCTCCCTCCAGCTCCAAAGCCACAGTCCGTAGCTTCAAGTCT	744
Qy	601	aagaa 605	
Db	745	AAAA 749	
RESULT	3		
LOCUS	BG720221	774 bp	mRNA
DEFINITION	60269227.F1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:4824760 5',		
KEYWORDS	EST.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 774)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapds@email.nih.gov		
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.		
	CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki		
	Toshiyuki and Piero Carninci (RIKEN)		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: LLM10736	row: m	column: 17
	High quality sequence stop: 769.		
FEATURES			
source	location/Qualifiers		
	1..774		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4824760"		
	/clone_id="NIH_MGC_97"		
	/lab_host="DH10B"		
	/note="Organ: testis; Vector: pBluescriptR (modified		
	pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcagag		
); Oligo-gt primed using primer 5'-TTTCTTTTCTTTTCTTTT-3',		
	size selected for average insert size 2.2 kb and		
	normalized to R01 5. This is a primary library enriched		
	for full-length clones and constructed using the		
	Cap-trapper method (Carninci, in preparation). Library		
	constructed by M. Brownstein (NIH/NHGRI, National		
	Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	225 a	185 c	198 g
ORIGIN			166 t
Query Match	55.8%;	Score 549;	DB 10; Length 774;
Best Local Similarity	98.1%;	Pred. No. 1.1e-143;	
Matches 609;	Conservative 0;	Mismatches 5;	Indels 7; Gaps 5;
Qy	1	atgtgtagaccagtggtgtgtgacctggaaagaaatcccccctagatgatgtacctaacc	60
Db	159	ATGTGTACACACAGGGGTGTGACCTGGAAGAAATCCCTCATGATGATGATGACCTAAAC	218
Qy	61	acatagaaattcaaaatcctcgactactacacagacatcatgttcttaagaagaccct	120

D	b	219	ACGATAGATTCAAAATCTCGCCTACTACACCAACAATCATGTCTTCAAGAGCACCCCT	278
O	y	121	gctcctcttcacccaagctgctgaaacaagaagtltgtccagagggcgcttgggaat	180
D	b	279	GCTCTCTTCTCACCACCAAGCTGCTGGAACAGAAGATTGTGCCAGAGGGGCGCTGGGGAAT	338
O	y	181	tgttcagcaaatagtcatagtgacagaggtytcataygccttcgcaaatccaatccagt	240
D	b	339	TGTTTCAGCAAAATGAGTCATGACAGAGGTGTCATGTGCCTTGAGAAAATTCCAAATCCAGT	398
O	y	241	gagaagggcataaaccttgycgaagaagaagctttcttgaaaagcatlcttbgagtagt	300
D	b	339	GAGAGGGCCATTAACTTGCGCAGAGAAAAAGTCTTGTGAAAAGCATCTTTGGAGTAGTG	458
O	y	301	gagaaggaagaattcgcacagacacgccttcgcaagtlctctgctcaggytcaaaagacgtg	360
D	b	459	GAGAGGAAGAATTCGCCAGAGGACGCGCTGCCAAGTCTCTGCTCAGAGGTCAAAAGACGTTG	518
O	y	361	gaataccaagattcgacacagccagtcagtgyltccagglytclttctaacytggagcagtc	420
D	b	519	GAATPCCAAGATTCGCCACAGCCAGCAGAGTGTCACAGTGTCTTTCTTAAGGTGAGCAGATGC	578
O	y	421	-ttgagagatbaagctfygagcccccaagatattccaatlgtccaaaccgagtagtcgaat	479
D	b	579	TTTGGAGATGAGGTGAGTGGACCCCAAAAGTCAATTTCCATTTGCCAACCGATAGCTGAAT	638
O	y	480	tgttatctccttgaccacacacacagcagaccagcagagagcttcaagtcacaaagat	539
D	b	639	TGTTTACCTCTGCG--ACACCAAGGCCA-CCAGGACGAGGCTTCAGTCCAAAGAGAT	695
O	y	540	tttgttaactgaaggctctctctcttcacagctcaa-ggccacgtgcctglaagcttaagt	598
D	b	696	TTTTTAACTGAGGCTCTCTCTTCCACTCACTCAAGGCGACAGCTGCCTGTAG--TTCAGTT	753
O	y	599	ctaagaagaatgaagaagaac	619
D	b	754	CTAAGACAGATGAAGAAGAAC	774
<hr/>				
RESULT 4				
A	KO16997	LOCUS	AKO16997	1535 bp mRNA linear HTC 19-JAN-2002
A	KO16997	DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:493430L01:BCL2-like apoptosis inhibitor containing protein, full insert sequence.	
A	KO16997	VERSION	AKO16997.1	GI:12856033
A	KO16997	KEYWORDS	HTC; CAP trapper.	
A	KO16997	SOURCE	Mus musculus (strain:C57BL/6J) adult male testis cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library clone:493430L01.	
<hr/>				
O	R	ORGANISM	Mus musculus	
O	R	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
O	R	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
R	AUTHORS	REFERENCE	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Hayashizaki,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	
R	AUTHORS	REFERENCE	High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)	
R	AUTHORS	REFERENCE	99279253	
R	AUTHORS	REFERENCE	10349636	
R	AUTHORS	REFERENCE	2 (sites)	
R	AUTHORS	REFERENCE	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Hayashizaki,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	
R	AUTHORS	REFERENCE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	
R	AUTHORS	REFERENCE	20499374	
R	AUTHORS	REFERENCE	11042159	
R	AUTHORS	REFERENCE	3 (sites)	
R	AUTHORS	REFERENCE	Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,U., Nishi,K., Kitsumai,T., Tashiro,H., Itoh,M., Suni,I., Ishii,Y., Nakamura,S., Hazema,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,	

QY	886	tttggcacaagaattacccttgaagaagacattcttcgcattgatccagcacacagctgtagt	945
Db	1137	TTTCGGACCACTGACTACCTTAAGACAGACTTCTCCCTGGCTTCACACAGATGGCGATGG	1196
QY	946	gaaaaatactctgggataltcacatgaagaagttagactga	984
Db	1197	GAAAAAATACCTTGGGATCTGCATCATGACAGATGAGACTGA	1235
RESULT	5		
LOCUS	AK008682		
DEFINITION	AK008682	1541 bp mRNA linear HTC 19-JAN-2002	
ACCESSION	AK008682		
VERSION	AK008682.1	GI:12843025	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male stomach cDNA to mRNA, clone:1210008009.		
ORGANISM	Mus musculus		
REFERENCE	1 (sites)	1 (sites)	
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2 (sites)		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3 (sites)		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,T., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,K., Inoue,K., Togawa,Y., Tanaka,M., Ohara,E., Watabiki,M., Oyazaki,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsui,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multipipillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE	4 (sites)		
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
MEDLINE	5 (bases 1 to 1541)		
PUBMED	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Nunazaki,R., Oho,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaoka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-JUL-2000)		

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAACAAGATCCAGACGCTCTTTTTCCTTTTTTTTAA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 20.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGCTTTCGAGTTAATTAAATTATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES

source	Location/Qualifiers 1..1541 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGD:MGI:1898348" /db_xref="taxon:10090" /clone="221008009" /sex="male" /tissue_type="stomach" /clone_id="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" 174..1160 /note="evidence:NAS hypothetical protein putative" /codon_start=1 /protein_id="BAB25830.1" /db_xref="gi:12843026" /translation="MCSYSDLDLPIEDDDPNSIEFKILAFARHHVFNTPAVFS PKLSRTISQKALGTSTDSMVQSYSPDRSPSEKNISIGKKSRITLFPAEEK EGLPSSKEPTIAQCPOGPFPYERSGFNMQMPSSLSEVDRLSEYVDSRYVALANR VAELVSWPPDVTHSQGSGLKERVSILTFREGEPCDSNKGEDDIISKIVELLK FGSGQLGEIKRKDALMSFFQDGSLSTFFKITDLFLMDVDTRGSEEVKANGFAALA IDAKLTALAIINHPMMNLRTGGTYLRFEYFSPWVOONGWEKIIGISHVEVD"
CDS	BASE COUNT 418 a 383 c 412 g 328 t ORIGIN

```

Query Match          53.6%; Score 527; DB 11; Length 1541;
Best Local Similarity 74.3%; Pred. No. 2,4e-137;
Matches 742; Conservative 0; Mismatches 230; Indels 27; Gaps 5;
```

OY	1 atgtctagaccacagtgggtgtgaccttggaagaatcccccatgatgatgataactaac 60
Db	174 ATGTCAAGCACACAGTGATGATACCCTCGAAGACATTCCTCCGTGAGAGATGATATCCAAC 233
OY	61 accatagaattaaaatcctgcctcactcacaccagaacatcatcgatttcgaagaccct 120
Db	234 AGCATAGAGATTAAAATCCTGCCTCTTTCAGGCCAGACACCATGATGCTTCAGAACACCCG 293
OY	121 gtctctctcccaaaagctgctggaacaagaagtlbtcccagagaggccttgsgaat 180
Db	294 GCTGTCTTCTGCCCAAGGTCCTCCAGAACAAAGAGTCTGTCCACAGAACCCCTGGGGAAT 353
OY	181 tgttcagcaaatgagtcatgatgacagagtgltcatgygaccttgcaaaattccaatcagt 240
Db	354 TGGT--CAACTGATTCCTCGACACAGATATCATGTGCTTCGACAGAGGTTCCCTTCACAC 410
OY	241 gagaaagccataaacttggcagaaaaaagtcttccttggaagacattcttggagtagag 300
Db	411 GAAGAAGACATAGCTTGCGCAGAGAGAGAGTCTTCTTGAGAACATCTTCAAGGGGGCC 470

QY	301	gagc---aggaaatctgcagacagccgcgaagctctctcgaagttc-----	349
Db	471	GAGAGGAGGAGAGGCCCTGCCGAGCTCCCAAGAAGATCCGAGCTCAGGTCCTCAGGCG	530
QY	350	---aaagacgttgaaataccaagattcgacagccagcagtgctcagtgcttct	405
Db	531	CCCTTCCCGGTAGAGCGCGCAGAGTGGCTTCCCAACACGACGACTCGGCCCAAGTCTTGAGC	590
QY	406	aacgttgagcagcgtcttgagacatgaagctcgtggaaccccaaatcaattccattgcgaac	465
Db	591	AGTGTGAGACAGCCCTCGGAGAGTGAAGTTGTGATTCCAAAGTGGCTTATTGGCAAC	650
QY	466	cgaatgcgtgaattgtttatctccgtgcacccacacaagcgaaacccagcgaagcgttc	525
Db	651	AGAGTGGCTGAATTTGTTTACTCTCTCGGCCACACAGATGTCATCCACACCAGGAGGA	710
QY	526	aagtcacaagagatllttgtlaactgagcgtctcccttcagctccaagccacgctgcct	585
Db	711	AGCAAGCTCAAAAGAGAGGCTCTGGAGATTTCTGTACTCAGGTTTGAAGAC---CTT	765
QY	586	gtacgttcaagttctaaagaagatlgaaagaacaatactagccaaatctgttgacgtg	645
Db	766	GTGACCTTAG---AATAAAGATGGAGGAGACCAATATATAGCAAGATTGTTGAGCT	821
QY	646	ctggaataatctcaggaagatcagtttggaagaagcgtgaaagaataagccttgatgggc	705
Db	822	CTGGAATTTCTCGGGGGTACGATTGGGGAAGAGATTAAGAAACAAACAGCGTTTGATGAGC	881
QY	706	cacttcaggaatggtcgtctctaactcgttttcaagaacatacagaaccagtcctaatg	765
Db	882	AGCTTCCAGGACGGGGCTGTCTACTCAACGTTCAAGACATCAAGACAGCTGTTCTGAGG	941
QY	766	ggtgttgacccccaaggaggagaatacagaagttcaaaagctcagggtcttaagctgccttga	825
Db	942	GACCTGGACACACAGAGAGATCATAGAGGTCTCGGGGCTTCAAGGCTGCGCTTGGCA	1001
QY	826	ataagctgcacgcgaacgctcacaagcatttgacaacaaccgcgtatgaacagcgtcccgagc	885
Db	1002	ATAGACCCCATCCGCAAGCTCACGGGCATGTGACACACCACCAATGAATGAATGCTGGGC	106
QY	886	tttgaccaagaactcgtgaagaagactctcgccabtgatccagaacacggttgatgg	945
Db	1062	TTCGGGACCAAGTACCTTAAGAGAGTACTTCTCCCGCTGGGTTACAGCAAGATGGCGATGG	112
QY	946	gaaaaaatacttgatatacatatgaagaagttagactga	984
Db	1122	GAAAAAATACTTGAGATCTACATGAAGAAGTGAAGCTGA	1160

RESULT	6
AK018579	
LOCUS	
DEFINITION	1578 bp mRNA linear HTC 19-JAN-2002
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
PubMed	
REFERENCE	
AUTHORS	

TITLE	Normaliztion and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3 (sites)
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carrinci,P., Konno,H., Atiyama,T., Nishi,K., Kitsuaki,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,I., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsumura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5 (bases 1 to 1578)
AUTHORS	Aachis,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bull,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Itawa,M., Kesukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishl,K., Nomura,K., Numataki,R., Ono,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schirml,L., Shibata,K., Shibata,Y., Shigaawa,A., Shiraki,T., Sobabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toyota,T., Yamamura,T., Yamataka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submissions
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGACGAATCCAGTAAATTAATTAATTCCTCCCCCCC 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot -10.0 and subtraction to Rot =185.2. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGACGAATTCCTCAGTTAATTAATTAATTCCTCCCCCCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluescript KS(+) after bulk excision from lambdaBla FLIC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.
FEATURES	Location/Qualifiers 1..1578 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="MGI:MGI:1894816" /db_xref="taxon:10090" /clone="9030625M01" /sex="male" /tissue_type="colon"

CDS

/clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 216..1202
 /note="evidence:NAS
 hypothetical protein
 putative"

/codon_start=-1
 /protein_id="Bab31290.1"
 /db_xref="GI:12858358"
 /translation="MSTSYVDLEDIEDDDPNISIEFKILAFYARHHVFKMPAVES
 PLISRTSRISOKALGTMTSDMQVSLPCRSPESEKISLGKSSMRTEPRVAKE
 EGIPLSPKREIDVAGPOGPFVEVERSGFNHNPRLSLEVDLSEVDSKYACIANR
 VAEIVSWPPDVIHSQGSGLKERVSEILFYREPGDSDNKDEDDIISKIVELK
 FSGDOLGREIKDKDRLNSFODGLSYSTFKITDLEFDVDTRGSEVAKRGFKALA
 IDAIKLTALDNPMLNMLGFGFKYLRREYFSPVQONGCWKEIILGISHEVD"

BASE COUNT 427 a 393 c 420 g 338 t
 ORIGIN

Query Match 53.6%; Score 527; DB 11; Length 1578;

Best Local Similarity 74.3%; Pred. No. 2.4e-137;
 Matches 742; Conservative 0; Mismatches 230; Indels 27; Gaps 5;

OY 1 agtgtgaagcaagctgggtgtgtgacctggaagaatccccctagatgatgatacctaac 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 216 ATGTGAGCAGCACAGTGTATGACCTGGAAACATTCCTCGAGAGATGATCCAAAC 275
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 61 accatagaattcaaaacccctcgactacacacagacatatgtcttcaaggaccacct 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 276 AGCATAGAGTTCACAAATCTTGCCCTTACCCAGACACCATGTCTTCAGAACACCCCG 335
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 121 gctctctctaccacaaagctgtctgagacaaagattgtccagagagggcttgaggaat 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 336 GCTGTCTTTCGCCAGCTCTCCAGAACAGAGTGTGTCCAGAACACCCCTGGGAGCT 395
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 181 ttgtcagcaaatgaatcatatgacagaggtgtcatgtgcttgcagaatcccaatcagat 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 396 TGGT---CAACTATTCTCGACACAGATATCATGTGCTTGAGAGGTTCCCTCCACAG 452
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 241 gagaagcgtaaaccttgcaagaagaagctctcttggaagacacctcttggaagtagtg 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 453 GAAAGAAACATCAGCTTGGGCAAGAGAGCTTCTTGAGAACACTCTTCCAGGTTGGCC 512
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 301 gaga---aggaagattcgacagacgctgcacagctctgtctcagagtc----- 349
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 513 GAGAGAGAGAGAGGCGCGGAGCTCCCAAGAGAGATCCAGCTCAGGGTCTCTCAGGGC 572
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 350 ----aaagaagcttggaatccaagattcgacagcagcagcagtggtccaggtgtcttct 405
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 573 CCCTTCCCGGTAGAGCGGAGAGTGGCTTCCACAAACGACACTGGCCCGGCTCTGAGC 632
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 406 aagctgagagcagctgtgagcagatgaaagctgtgaccccaagatcttcattgccaac 465
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 633 AGGTGTGAGAGAGCGCCCTGGAGAGTGAAGTTTGGATTCCAAAGTGGCTTGTATTGGCAAC 692
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 466 cgaagtagtgaatattgttattctctgacacacacacagcagccagagcaggaagcttc 525
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 693 AGAGTGGCTGAATTTGTTTACTCTGCGCACACAGATGTCATCCACAGCCAGGAGGA 752
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 526 aagtcacaagaagatttttgaactgagaggtctctccctcagctccaagcagctgacct 585
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 753 AGCAAGCTCAAGAGAGGCTCTCGAGATTGTGATCTCAGGTTGAAAGAC-----CTT 807
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 586 gtatgctcaagcttcaagaagaatgaaagaacaatacttcgcaaaatttgtagctg 645
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 808 GTTACTCTTAAG-----AATTAAGATGGAGAACCAATTAATAGCAAGATTGTTGAGCTG 863
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 646 ctgaatatctcaggaagctcagcttggaagaagaagctgaaagaatagaagctttagatggc 705
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 864 CTGAATAATTCTCGGGGAGATCAAGTTGGAGAGAGATTAAGAAAGACAGGCTTGTGATGAG 923
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 706 caattcaggaatgggtcgtctactcgttttcaagaacatcagaagcagctgctctaag 765
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 924 AGCTTCAGGAGGCGGTGCTCTACTCAAGCTTCAGACCATCAAGACCTGTTCTGAGG 983
 OY 766 ggtgtgagcccccaggggaagaatcatagagatcagaagctcagaagctttagagcttcgtta 825
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 984 GACGTGACACCAAGAGAGATCATAGAGTCAAGCTCGGGGCTTCAAGGCTGCGCTTGCA 1043
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 826 atagacgtcagcgccaagctcacaagctatctgacacacacccagatgaacaggtctctggc 885
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1044 ATGACAGCGCATGCGCAAGCTCAGCGGCATTCAGACACACCATGATGATAGATGCTGGCC 1103
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 886 ttgtgacccaagctcagtaagaagaactctcgcatgtagtccaagcagcagctgtagtg 945
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1104 TTGGGACCAACTACTAGAGAGTCTCTCCCTCGGGCTTACGAGATGGCGGATGG 1163
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 OY 946 gaaaataactctgagatcatcatgaaagaagtagactga 984
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1164 GAAATAATCTTGGGATCTCAGATGAAGATGACTGA 1202
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7

AK016670
 LOCUS 2222 bp mRNA linear HTC 19-JAN-2002

DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933405K19:BCL2-like apoptosis inhibitor containing protein, full insert sequence.

ACCESSION AK016670
 VERSION 1 GI:12855533

KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (Strain:C57BL/6J) adult male testis cDNA to mRNA, clone:4933405K19.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning

TITLE

Meth. Enzymol. 303, 19-44 (1999)
 99279253

JOURNAL

MEDLINE 10349636
 PUBMED

REFERENCE

2 (sites)
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

TITLE

20499374
 MEDLINE 11042159
 PUBMED

REFERENCE

3 (sites)
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishii, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

TITLE

20530913
 JOURNAL 11076861
 MEDLINE
 PUBMED

REFERENCE

4 (sites)
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

TITLE

5 (bases 1 to 2222)
 Adecchi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arai, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,

JOURNAL

REFERENCE
 AUTHORS

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 715)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: L14M11560 row: 1 column: 20 High quality sequence stop: 715.
FEATURES	Location/Qualifiers
source	1..715 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:522875" /clone_1ib="NIH-MGC_120" /lab_host="DH10B" note="Organ: pooled pancreas and spleen; Vector: pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH-MGC Library."
BASE COUNT	212 a 177 c 171 g 155 t
ORIGIN	
Query Match	51.5%; Score 506.6; DB 10; Length 715;
Best Local Similarity	99.2%; Pred. No. 9.9e-132;
Matches 509; Conservative	0; Mismatches 4; Indels 0; Gaps 0.
1 atgtgttagcaccagctgggtgtgtgacctgtggaagaatcccccctagatgatgtacctaac 60	
203 ATGTGTAGCACACAGTGGGTGTGACTGTGGAAGAAATCCCCCTGATGATGATGACCTAAGC 262	
61 accatagaattcaaaatccctgcctactacaccagacatcatgtcttcaagaacccct 120	
263 ACCATAGAATTCAAATCCTGCCTTACTACACACACATCATGCTTCAAGAGCACCCCT 322	
121 gctcctcttcaccaaagctgctgtgagaacaaagtgtgtcccaagaaggcctgtggagt 180	
323 GCTCTCTTCTCACCAAGCTGCTGGAACACAAATTTGTCCCAAGGGGCTTG66GAT 382	
181 tgttcagaagaatgtcatgtgacaagaaggtgtgcatgtgacctgcagaaatcccaatccagt 240	
383 TGTTCAGCAATGATCATGTGACAGAGGTGTATGGCTTGACGAATTCCTCAATCCACT 442	
241 gagaagccataaaccctgtgcaagaagaagtcctcttggaaagcatctcttggagtagtg 300	
443 GAGAAGGCCATTAACCTTGGCAAGAAAAAGTCTTGTGGAAAGCATTTCTTGGAGTAGTG 502	
301 gagaaggaagatctgcagagacagcgtctgccaagtgctcgtcgaaggttcaagaagacttg 360	
503 GAGAAGGAAGATTCCAGAGGACAGCGCTGCAGAGGCTCTGCTCAGGGTCAAGAGAGCTTG 562	
361 gaataccaagaattgcacagccagcagtgctgcaggtgtctcttcaacgtggagcaatgc 420	
563 GAATTCACCAAGATTGCCACACGCCAGCAGTGTGCTCAGGTGTCTTTCAAGCTGGACAGTGC 622	
421 ttggagcaagaagctgtggaaccccaagtcattccattgccaaccgagtagctgaatt 480	

D6	623	TTGAGCAGTGAAGCTGTGGACCCCAATCATTTTCATTGCCAACCGTAGCTAATC	682
OY	481	gttattcttgccaccacccaagaagcaccag	513
Db	683	GTTACTCTCTGGACACCAACAAGCACCCAG	715
RESULT	9		
LOCUS	Bg717835	672 bp	mRNA linear EST 08-MAY-2001
DEFINITION	60269385F1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:4826292 5',		
ACCESION	Bg717835	Homo sapiens	
VERSION	Bg717835.1	GI:13997022	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteria; Primates; Catarrhini; Hominoidea; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs@mail.nih.gov Tissue Procurement: Mikhail Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shireki Toshiyuki and Piero Carninci (RIKEN) CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov Plate: LRAM10740 row: m column: 13 High quality sequence stop: 648. Location/Qualifiers 1..672 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4826292" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgaggg...); Oligo-dT primed using primer 5'-tttttttttttttttttttttVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT	197 a	162 c	171 g
ORIGIN	142 t		
Query Match	50.5%; Score 497.2; DB 10; Length 672;		
Best Local Similarity	99.2%; Pred. No. 4.3e-129;		
Matches	510; Conservative	0; Mismatches	3; Indels
			1; Gaps
OY	1	atgtgttagacacagtgtgagtgtaacctggaagaatccocctagatgatgatgaactaac	60
Db	159	ATGTGTACACACAGGGGTTGACTTGGAAGAATCCCCCTGATGATGATGACTTAAC	218
OY	61	accatagaattcaaatctctgcctactaacaccaacaatcatgctttaagaagcacccct	120
Db	219	ACCATAGAAATTCATAATCTCGCTACACACCAACATCATGTCTCAAGACACCCCT	278
OY	121	gtctcttcctacccaagctgtctggaacaagaatttgtcccagaaggcctgtggagat	180
Db	279	GCTCTCTTCTGACCAACAGCTGCTGAGACAAACAATTTTGTCCCAAGGGGCTGGGGAAT	338
OY	181	tgttcagaanaatgagtcatggaacagagtgltcatgagccttgacagaatcccaatccag	240

```

Db 339 TGTTCGCAAAATGATCATGACAGAGGTGTCATGCGCTTCGAGAAATCCCAATCCACT 398
OY 241 ggaagaagccataaaccttgcgaagaagaagctctcttggaagaacattcttggatgaag 300
Db 399 GAGAAGGCGCATTAACCTTGCCAGAAAAGCTCTTGGAAAGCATCTTTGGAGTAGTG 458
OY 301 gagaagaagatctgcgaagcaagcctgcgaaggtctctctgcaggtgcaaaagacgttg 360
Db 459 GAGAAGAGAGATTCGACAGACAGCGCTGCGCAAGGTCTGTGCTCAAGGTCAAGAGACGTTG 518
OY 361 gaataccaa-gatctgcacagcagcagcagtgctcagtgcttcttaacgttgagacgtg 419
Db 519 GAATACCAACGATTCGACAGACAGCAGGTGCTCAAGGTCTTCTTAACGTGAGACAGTG 578
OY 420 ctggagagcattgaagctgttgagaccgaagatcttcattccatgcgaacgagtagctgaat 479
Db 579 CTGGAGCATGAAGCTGTGGACCCCAAGTCATTTCCATTGCCAACCGAGTAGCTGAAT 638
OY 480 tgttatctctggccacaccacacaagcgaccag 513
Db 639 TGTATTCCTCTGTCACACACACAGGAAACAG 672

```

```

RESULT 10
BI908814 703 bp mRNA linear EST 16-OCT-2001
LOCUS 60306627F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5215218 5',
DEFINITION MRNA sequence.
ACCESSION BI908814
VERSION BI908814.1 GI:16171889
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 703)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM11540 row: J column: 19
High quality sequence stop: 701.
Location/Qualifiers
1..703

```

```

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5215218"
/clone_1ib="NIH_MGC_118"
/tissue_type="Leukocyte"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6, Site.1: NotI, Site.2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH_MGC Library."
BASE COUNT 208 a 173 c 170 g 152 t
ORIGIN
Query Match 42.5%; Score 418.4; DB 10; Length 703;

```

Best Local Similarity 98.3%; Pred. No. 7,le-107;
Matches 507; Conservative 0; Mismatches 1; Indels 8; Gaps 8;

```

OY 1 atggtgacaccagaggggtgactctggaagaatccccctagatgatgtgacctaacc 60
Db 190 ATGTGTACACCAAGTGGGTGTGACTGTGAGAAATCCCTTATATATGATGACCTTAAC 249
OY 61 accataagattcaaaatctcgcctactacacacagac-atcattcttcaagaaccccc 119
Db 250 ACCATAGAAATTCAAATCTCTGCGCTACTACACAGACTATCATGTCTTCAAGAGCACCCC 309
OY 120 tgcctctctcaccacaagctgtctgagaacaagaattgtccagaagggtctgggaa 179
Db 310 TGCTCTCTTCTCACCAAGAGCTGCTGAGAAACAAGATTGTCTCCAGAGGGGCTGGGAA 369
OY 180 ttgttcgcaaatgattatgacagaggtgtcattgaccttggaaatctccattccacag 239
Db 370 TTGTTTCACCAATGATGATGACAGAGGTGTATGAGCTTGTGCAAGAAATCCCAATCCAG 429
OY 240 tgaagaagccataaaccttgcgaagaa-aaagtctcttggaaagattctt-tgagta 297
Db 430 TGAGAAGGCGCATTAACCTTGCGAAGACAAAGTCTTGTGGAAGCATTTGTGAGTA 489
OY 298 gtggaagaagaagattcg-cagagcaagcctgcgaaggtctctctcaggggtcaaaagac 356
Db 490 GTGGAAGAAGGAAGATTCCGACAGACAGCGCTGCCAAGGTCTGCTCAGAGTCAAGAGAC 549
OY 357 gtgtgaatcacaagatttgcacagc-cagcaatgtgtccaggtgtcttcttaa-cgtgag 414
Db 550 GTTGAATACCAAGATTGTGACAGCGCACAGTGTCTCAGGTCTTCTTAACCGTGGAG 609
OY 415 cagtccttgagcatgaagctgttgaccaccaagatcttcattccatgcgaacgagtagct 474
Db 610 CAGTGTCTTGAGCATGAAGCTGTGAGACCCCAAGTCTTCCATTGCCAACCGAGTAGCT 669
OY 475 gaattgttatctctggccacaccacaagcgacc 510
Db 670 GAAA-TGGTATTCCTCTGCGCA-CACACACAGCAGAC 703

```

```

RESULT 11
AW000827/c 426 bp mRNA linear EST 09-MAR-2000
LOCUS W445410.x1 Soares.Dieckgraefe.colon_NHCD Homo sapiens cDNA clone
DEFINITION IMAGE:2522995 3', mRNA sequence.
ACCESSION AW000827
VERSION AW000827.1 GI:5847665
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 426)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 634 Std Error: 0.00
Seq primer: -40UP from Glpco
High quality sequence stop: 419.
Location/Qualifiers
1..426

```

```

FEATURES
source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2522995"
/clone_1ib="Soares.Dieckgraefe.colon_NHCD"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"

```

/note="Organ: colon; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5', TGTACCAATCTGAAGTGGAGCGCGCCGCTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieckgrbm.wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 91 a 108 c 88 g 139 t

ORIGIN

Query Match 38.2%; Score 376; DB 9; Length 426;
Best Local Similarity 99.3%; Pred. No. 5.2e-95;
Matches 399; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

584 ctgtagcttaagcttaagaagaatgaagaacaatactag-ccaaattgttgg 642
|||||
426 CTGTAGCTTAAGCTTCTAGAGAAAGTGAAGAACAAATCTAGCCAAATTTGTGAG 367
643 ctgctgaatattcagagagatcagcttggaaagaagctgaagaataagcttgaatg 702
|||||
366 CTGCTGAATATTTCAGAGATCAGTGTGAAGAAAGC--AAGAAGATTAAGCTTTGATG 309
703 ggcacactcagaagatggctgtcctactcgttcttaagaacatcacagaaccagttccta 762
|||||
308 GGCACACTCCAGAGATGGGCTCTCTACTCTGTTTCAAGACCATCACAGACGATCTCTA 249
763 atgggtgttgagaccccgaggagaaatcagaagctcaagggcttaagtgccctt 822
|||||
248 ATGGGTGTGAGACCCAGGAGGAATCAGAGGTCAAGGCTCAGGGCTTTAAGGCTGCCCTT 189
823 gtaataagctcagagcgaagctcagactatgtacaacaccgcggaatgaagaagttcctg 882
|||||
188 GTATATGACGTACAGGCCCAACTCTACACTATTGACAAACCCGATGAACAGGTCCTG 129
883 ggccttggcaccagatcctcgtaaagaagactctgcgcattgcatcagaagcaggttga 942
|||||
128 GGCCTTGGCACCAAGTACTCTGAAGAGAACTTCTGCGCATGATCCAGACGACGCTGA 69
943 tgggaataaatacttggtgatatacatalgaagaagtagactga 984
|||||
68 TGGGAAAAAATCTTGGGATATCATGAAAGATGAGACTGA 27

RESULT 12

LOCUS B1652080

DEFINITION 603298866F1 NIH_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5339464 5',

ACCESSION B1652080

VERSION B1652080.1 GI:15566316

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 676)

NIH-MGC <http://mgs.nci.nih.gov/>.

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaps@email.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1M1862 row: c column: 17

High quality sequence stop: 676.

FEATURES

Source

1..676

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="IMAGE:5339464"

/clone_lib="NIH_CGAP_Mam3"

/tissue_type="tumor, gross tissue"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert 2 kb. Library constructed by Life Technologies, catalog #12017-018. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference

for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999). Note: this is a NIH CGAP Library."

BASE COUNT

194 a 156 c 189 g 137 t

ORIGIN

Query Match 36.5%; Score 359; DB 10; Length 676;
Best Local Similarity 78.9%; Pred. No. 3.9e-90;
Matches 482; Conservative 0; Mismatches 115; Indels 14; Gaps 4;

376 ccaagcagcagtggtccagggtgtcttcttaagtggaagcagtgcttggagcagaact 435
|||||
16 CACAACACACACTGCGCCAGGCTCTGAGCAGTGTGAGCAGCGCTGGAAGTAAAGTT 75
436 gtggaccccaagatcattccatggccacccagatagctgaattgtttatctctggca 495
|||||
76 GTGGAATCCAAAGTGGCTTGTATTGCCAACAAGTGGCTGAATTTACTCTCGGCCA 135
495 ccaccacaagcgca--ccagycagagaggtctcaagttccaaaagatttgttaactgag 553
|||||
136 CCACCAAGATGTCATCCACAGCCAGAGGAGAGCAAGCTCAAGAGA--GGGTCTCGAGA 192
554 gttctctcttccagctcccaagcggtgcctgttagcttaagttcagaagtaagttgag 613
|||||
193 TTCTGTACTTCAAGGTTGAAGGAC---CTTGTGACTCTTAAG---AATTAAGTGTGAG 243
614 aagaacaatactagcacaataattgttagctgtgaaatattcagagatcaattggaaa 673
|||||
244 AAGACCAATTAATAACCAAGATGTTAGCTGCTGAATTTCTCGGGGATCAGTTGGGA 303
674 gaaagctgaagaagaataaggtcttgaatggccaacttccaagatgggtgtcctactctg 733
|||||
304 GAGAGATTAAGAAAGACAAGGCTTTGATGAGACACTTCCAGAGCGGGCTGTCTACTCA 363
734 ttctaaagccatcaagaccaggtctcttaatgtgtgtgagcccaagggtgaataagagg 793
|||||
364 CGTTCAAGACCAATCAGACACTTGTCTGAGGAGCGTGGACACACAGAGAAATACAGAG 423
794 tcaagctcaaggtcttaagctgtgccccttgaatagaagtcagagcgaagtcacagcta 853
|||||
424 TCAAGCTCGGGCTTCAAGGCTGCGCTTCGAATAGACGCCATGGCCAGCTACAGGGCA 483
854 ttgacaacaccacccgatgaagaaggtctctgttgacccaagatlaactgaagaagaact 913
|||||
484 TCGACCAACCAACCAATAGATAGATCGTGGGCTTCGGGACCAAGTACTAAGAGTACT 543
914 ttctgcacatggatccagcagcagcaggtgagatggggaataaataactgtggatcacatgag 973
|||||
544 TCTCCCTGCGGTTTCAGCAGAAATGGGATGGGAAAAAATATCTGGGATCTCATATGAG 603
974 aagtagactga 984
|||||
604 AAGTAGACTGA 614

RESULT	13
LOCUS	AQJ31601 422 bp DNA linear GSS 23-SEP-1998
DEFINITION	H5_3049.AL.D07_MR.CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3049 COL=13 Row=G, DNA sequence.
ACCESSION	AQJ31601
VERSION	AQJ31601.1 GI:3508767
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 422) Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
LINE	99380589
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel.: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Sequence Tagged Connector Plate: 3049 row: G column: 13 Class: BAC ends High quality sequence stop: 422. Location/Qualifiers 1..422 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="plate:3049 Col=13 Row=G" /clone_lib="CIT Approved Human Genomic Sperm Library D" /sex="male" /note="Organ: sperm; Vector: pBelOBAC1; BAC Clones In E-Coli DH10B"
BASE COUNT	110 a 101 c 104 g 105 t 2 others
ORIGIN	
Query Match	36.3%; Score 356.8; DB 12; Length 422;
Best Local Similarity	95.1%; Pred. No. 1.3e-89;
Matches 367; Conservative	0; Mismatches 19; Indels 0; Gaps 0;
D68	aattcaaaatctcgccactacaccagacatcatgcttccaaggacaccctgtcct 127
D36	attttcaaatcttcggccrctacacacacacatgatgctttcaagagcacacctgtcct 95
OY128	tctaccaaaagtctgtgagaacaagaatttgctcccgaaggggcctggggaatttaag 187
DB96	tcttacccttactctgttgaaacaaagaaattgtgccacagggcctgcggaatttgtag 155
OY188	caaatgaatctgacaagaagtgatcatgaccttgtagagaattccccaacagttgagaag 247
DB156	cattatgaactcatgacacagagtgctcatgtgaccttgcagaaaatttcccaattccactgagaaag 215
OY248	ccataaactctgycagaagaaaaagctcttcttgynaagcattcttggagtagtggagaag 307
DB216	ccatpaaaccttgccgaagaaaaagcttcttggaagacatcttggagtagtggagaag 275
OY308	aagattcgcaagaacagcgtctccaaagtctctgtctcaaggttcaagaagcgtttgataac 367
DB276	aagattcgcaagaacagcgtctccaaagctctgtctccaaaggttcaagaagcgtttgataac 335
OY368	aagattcgcaagaacagcgtctccaaaggtctctgtctcaaggttcaagaagcgtttgataac 427
DB336	aagattcgcaagaacagcgtctccaaaggtctctgtctcaaggttcaagaagcgtttgataac 395
OY428	atgaagctgtggaccccaagtcatt 453

Db	396	ATGACGTCATGCACTCGGATCTTT	421
RESULT 14			
LOCUS	BB013694	984 bp	mRNA linear EST 18-OCT-2001
DEFINITION	BB013694 RIKEN full-length enriched, adult male testis (DH10b) Mus musculus cDNA clone 430470009 3', mRNA sequence.		
ACCESSION	BB013694		
KEYWORDS	BB013694.2 GI:16257339		
ORGANISM	EST. house mouse. Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus. 1 (bases 1 to 984)		
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.		
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)		
JOURNAL	Unpublished (2001)		
COMMENT	On Jun 1, 2000 this sequence version replaced gi:8136849. Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenhiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) Wagii,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Location/Qualifiers 1. 984 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="4930470009" /clone_1lb="RIKEN full-length enriched, adult male testis (DH10b)"		

Db 292 GTGAGAGGCCATTAACCTTGCAAGAAAAAGCTCTTGGAAAGCATTTGGAGTAG 351
 Oy 299 tggagaggaagaatcgcaagacacgcctgccaaggtctgctcaggtcaag 353
 |||||
 Db 352 TGGAGAGGAGATTCGAGAGCACGCTTGCCAAAGTCTTGTCTCAGGCTCAAG 406

Search completed: June 20, 2002, 04:19:16
 Job time: 6406 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 16:33:47 ; Search time 15.65 Seconds
(without alignments)
809.028 Million cell updates/sec

Title: US-09-771-961-2

Perfect score: 1709
Sequence: 1 MCGTSGCDLEIPDDDLN.....WIOHGGEMKILGISHEVD 327

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Indexed: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	128.5	7.5	239	1	BCLX_CHICK
2	119	7.0	233	1	BCLX_MOUSE
3	118	6.9	233	1	BCLX_RAT
4	117	6.8	233	1	BCLX_HUMAN
5	115	6.7	530	1	TDT_MOUSE
6	114	6.7	233	1	BCLX_PIG
7	105.5	6.2	901	1	SECA_HAFTN
8	99	5.8	787	1	YVDR_BACSU
9	97	5.7	537	1	SYK_AQUAE
10	96	5.6	2319	1	AKA6_HUMAN
11	94.5	5.5	192	1	BAXA_MOUSE
12	93	5.4	204	1	AR11_XENLA
13	93	5.4	859	1	PMS2_MOUSE
14	92.5	5.4	723	1	GGA3_HUMAN
15	92.5	5.4	945	1	RPN2_YEAST
16	92	5.4	880	1	DPO1_BACSU
17	91.5	5.4	520	1	TDT_BOVIN
18	91	5.3	875	1	Z1P1_YEAST
19	90.5	5.3	192	1	BAXA_RAT
20	90	5.3	433	1	KAR1_YEAST
21	89.5	5.2	218	1	BAXB_HUMAN
22	89.5	5.2	696	1	UVRG_DROME
23	89.5	5.2	2319	1	FAB_MOUSE
24	89	5.2	926	1	POOL_HAFTN
25	89	5.2	1722	1	RBB2_HUMAN
26	88.5	5.2	445	1	FLDL_VIRPA
27	88.5	5.2	2524	1	NOTC_XENLA
28	88.5	5.2	3027	1	POLG_PYEVI
29	88	5.1	196	1	TECI_YEAST
30	87.5	5.1	482	1	BAXA_HUMAN
31	87	5.1	192	1	TDT_MONDO
32	87	5.1	854	1	AKA1_RAT
33	87	5.1	947	1	TOP1_MYCLE

34	86.5	5.1	1165	1	SYV_AQUAE
35	86.5	5.1	2875	1	RPL7_TSMV1
36	86	5.0	537	1	IMDH_DROME
37	86	5.0	856	1	CLPB_HELPY
38	85.5	5.0	250	1	YC97_HUMAN
39	85.5	5.0	1158	1	R114_HUMAN
40	85	5.0	229	1	BCL2_BOVIN
41	85	5.0	239	1	BCL2_HUMAN
42	85	5.0	639	1	GGA1_HUMAN
43	85	5.0	702	1	DY12_ANTCR
44	84.5	4.9	192	1	BAXA_BOVIN
45	84.5	4.9	697	1	YE9C_SCHPO

ALIGNMENTS

RESULT	ID	STANDARD	PRT	229 AA
1	BCLX_CHICK			
BCLX_CHICK	007816; Q98908;			067411 aquifex aeo
AC	007816; Q98908;			P28976 tomato spot
DT	01-FEB-1995 (Rel. 31, Created)			O07152 drosophila
DT	01-NOV-1997 (Rel. 35, Last sequence update)			P71404 helicobacte
DT	16-OCT-2001 (Rel. 40, Last annotation update)			Q9Y345 homo sapien
DE	Apoptosis regulator Bcl-X.			P48552 homo sapien
GN	BCL2L1 OR BCLX OR BCL-X.			O02718 bos taurus
OS	Gallus gallus (Chicken).			P10415 homo sapien
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			O9UJ55 homo sapien
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			O16959 anthocidari
OC	Gallus.			O02703 bos taurus
ON	NCBI_TaxID=9031;			O13773 schizosacch
OX	[1]			
RP	SEQUENCE FROM N.A. (SHORT FORM).			
RX	MEDLINE=93364977; PubMed=8358789;			
RA	Boise L.H., Gonzalez-Garcia M., Postema C.E., Ding L., Lindsten T.,			
RA	Turka L.A., Mao X., Nunez G., Thompson C.B.;			
RT	"bcl-x, a bcl-2-related gene that functions as a dominant regulator			
RT	of apoptotic cell death.";			
RL	Cell 74:597-608(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A. (LONG FORM).			
RC	STRAIN=HUBBARD WHITE MOUNTAIN; TISSUE=Testis;			
RA	MEDLINE=97264485; PubMed=9110311;			
RT	Vilagrassa X., Mezquita C., Mezquita J.;			
RT	"differential expression of bcl-2 and bcl-x during chicken			
RT	spermatogenesis.";			
RL	Mol. Reprod. Dev. 47:26-29(1997).			
CC	-1- FUNCTION: DOMINANT REGULATOR OF APOPTOTIC CELL DEATH. THE LONG			
CC	FORM DISPLAYS CELL DEATH REPRESSOR ACTIVITY, WHEREAS THE SHORT			
CC	ISOFORM PROMOTES APOPTOSIS (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: MITOCHONDRIAL MEMBRANES AND PERINUCLEAR			
CC	ENVELOPE (BY SIMILARITY).			
CC	-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, A LONG FORM (SHOWN HERE) AND A			
CC	SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.			
CC	-1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ORGANS WITH LYMPHOID			
CC	DEVELOPMENT.			
CC	-1- DOMAIN: BH4 DOMAIN SEEMS TO BE INVOLVED IN THE ANTI-APOPTOTIC			
CC	FUNCTION. INTRACT BH1 AND BH2 DOMAINS ARE REQUIRED FOR ANTI-			
CC	APOPTOTIC ACTIVITY (BY SIMILARITY).			
CC	-1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 1 (BH1).			
CC	-1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 2 (BH2).			
CC	-1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 3 (BH3).			
CC	-1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 4 (BH4).			
CC	-1- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			

DR EMBL: U10100; AAA82172.1; -
 DR EMBL: U51279; AAC53460.1; -
 DR EMBL: U78031; AAB96881.1; -
 DR EMBL: U78030; AAB96881.1; JOINED.
 DR HSSP: P53563; IAF3.
 DR MGI: 88139; Bcl2l.
 DR InterPro: IPR002475; BCL2_family.
 DR InterPro: IPR003093; BH4.
 DR InterPro: IPR00712; Bcl_2.
 DR Pfam: PF00452; Bcl-2; 1.
 DR Pfam: PF02180; BH4; 1.
 DR SMART: SM00337; BCL; 1.
 DR SMART: SM00265; BH4; 1.
 DR PROSITE: PS0062; BCL2_FAMILY; 1.
 DR PROSITE: PS01080; BH1; 1.
 DR PROSITE: PS01258; BH2; 1.
 DR PROSITE: PS01259; BH3; 1.
 DR PROSITE: PS01260; BH4; 1; 1.
 DR PROSITE: PS0063; BH4_2; 1.
 DR Apoptosis; Mitochondrion; Alternative splicing; Transmembrane.
 FT DOMAIN 4 24 BH4.
 FT DOMAIN 86 100 BH3.
 FT DOMAIN 129 148 BH1.
 FT DOMAIN 180 195 BH2.
 FT TRANSMEM 210 226 POTENTIAL.
 FT VARSPLIC 126 188 MISSING (IN ISOFORM BCL-X(S)).
 FT VARSPLIC 189 233 DTFVLDLGNNAAESRKGQEFNRFGLMTVAGVLLGSL
 FT VARSPLIC 189 233 FSRK -> VRTPLVCPPLACVSLCEHP (IN ISOFORM
 BCL-X(BETA)).
 FT VARSPLIC 194 233 LYGNNAAESRKGQEFNRFGLMTVAGVLLGSLFSRK
 FT VARSPLIC 194 233 -> GHDCGKGSAGLTLOSEVTRH (IN ISOFORM BCL-
 X(DELTA-TM)).
 SO SEQUENCE 233 AA; 26132 MW; 24D2AC79887E072E CRC64;
 Query Match 7.0%; Score 119; DB 1; Length 233;
 Best Local Similarity 21.1%; Pred. No. 0.033;
 Matches 54; Conservative 37; Mismatches 91; Indels 74; Gaps 11;
 QY 77 SOSSEKAI-----NLGKKSSKAFGVVEKEDSQSTPAKVSAGQRTLEYQDSHQOM 130
 DB 2 SQSNRLVYDFLSYKLSQKQYVSWQSFQDVENR-----TEAPRETEAKERTPAININPMS 57
 QY 131 SRCLSNVEQCLEHBAVDPKYIISANRVAEIVSWPPQATGAGFKSEIEFTGISTPOL 190
 DB 58 -----HLADSPAV-----NGAT----- 69
 DB 191 OGHVPAVSSSKKDEEEDL-LAKIYELKYSGDLE---RLKKKKALMGHPDGLSYSVF 246
 DB 70 -GH-----SSLDAREYIPMAAVKQALREADEDELRVRRAFSDLTSLHTPGTAQSF 123
 QY 247 KTTIDVLM-GVDPGRSEVKAQGFKAALVLDVTAKLTAIDNHPNRRVLFSGTKYLKENF 305
 DB 124 EQGVNLEFRGVN--WRIYAFPSFGALCVESYDKEMOV---LYSRILASMATIYINDHL 178
 QY 306 SEWIOGHGMEKILGT 321
 DB 179 EPVIDENGWDFVDL 194
 RESULT 3
 BCLX_RAT STANDARD; PRT; 233 AA.
 AC P53563; Q62678; P70614; P70613; Q62836; Q64087; Q64128;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Apoptosis regulator Bcl-x.
 GN BCL2L1 OR BCL2L OR BCLX.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxID=10116;

RN [1]
 RC SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(S)).
 RP TISSUE=Brain;
 RA Michaelidis T.M.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Brain;
 RA Messeligh S.L., David G.L., Choi S., Velluona M., Hardwick J.M.;
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RC SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(BETA)).
 RP TISSUE=Thymus;
 RX MEDLINE=96278736; PubMed=8662675;
 RA Shirlawa N., Inohara N., Okada S., Yuzaki M., Shoji S.-I., Ohta S.;
 RT "An additional form of rat Bcl-x, Bcl-xbeta, generated by an
 RT unspliced RNA, promotes apoptosis in promyeloid cells.";
 RL J. Biol. Chem. 271:13256-13265(1996).
 RN [4]
 RC SEQUENCE FROM N.A. (ISOFORMS X(L) AND X(S)).
 RP STRAIN=SPRAGUE-DAWLEY; TISSUE=ovary;
 RX MEDLINE=95129487; PubMed=7828536;
 RA Tilly J.L., Tilly K.I., Kenton M.L., Johnson A.L.;
 RT "Expression of members of the bcl-2 gene family in the immature rat
 RT ovary: equine chorionic gonadotropin-mediated inhibition of granulosa
 RT cell apoptosis is associated with decreased bax and constitutive
 RT bcl-2 and bcl-xlong messenger ribonucleic acid levels.";
 RL Endocrinology 136:232-241(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=9801630; PubMed=9346936;
 RA Arltom M., Kunishima N., Inohara N., Ishibashi Y., Ohta S.,
 RA Morikawa K.;
 RT "Crystal structure of rat Bcl-xL. Implications for the function of
 RT the Bcl-2 protein family.";
 RL J. Biol. Chem. 272:27886-27892(1997).
 CC -1- FUNCTION: Potent inhibitor of cell death. Inhibits activation of
 CC caspases (By similarity). Appears to regulate cell death by
 CC blocking the voltage-dependent anion channel (VDAC) by binding
 CC to it and preventing the release of the caspase activator,
 CC cytochrome c, from the mitochondrial membrane. The Bcl-x(S) and
 CC Bcl-x(beta) isoforms promote apoptosis.
 CC -1- SUBUNIT: Bcl-x(L) forms heterodimers with BAX, BAK and Bcl-2 (By
 CC similarity). Heterodimerization with BAX does not seem to be
 CC required for anti-apoptotic activity (By similarity).
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MEMBRANES AND PERINUCLEAR
 CC ENVELOPE (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; BCL-X(L) (SHOWN HERE), BCL-X(S)
 CC AND BCL-X(BETA); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST TISSUES. BCL-X(BETA) IS
 CC SPECIFICALLY EXPRESSED IN CEREBELLUM, HEART, AND THYMUS. IN THE
 CC OVARY, THE PREDOMINANT FORM IS BCL-X(L), WITH A SMALL BUT
 CC DETECTABLE LEVEL OF BCL-X(S).
 CC -1- DOMAIN: The BH4 domain is required for anti-apoptotic activity.
 CC The BH1 and BH2 domains are required for both heterodimerization
 CC with other Bcl2 family members and for repression of cell death.
 CC -1- PTM: Proteolytically cleaved by caspases during apoptosis. The
 CC cleaved protein, lacking the BH4 domain, has pro-apoptotic
 CC activity (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 1 (BH1).
 CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 2 (BH2).
 CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 3 (BH3).
 CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 4 (BH4).
 CC -1- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; BCL-X(L) (SHOWN HERE), BCL-X(S)
CC AND BCL-X(BETA); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: BCL-X(S) IS EXPRESSED AT HIGH LEVELS IN CELLS
CC THAT UNDERGO A HIGH RATE OF TURNOVER, SUCH AS DEVELOPING
CC LYMPHOCYTES. IN CONTRAST, BCL-X(L) IS FOUND IN TISSUES CONTAINING
CC LONG-LIVED POSTMITOTIC CELLS, SUCH AS ADULT BRAIN.
CC -1- DOMAIN: The BH4 domain is required for anti-apoptotic activity.
CC The BH1 and BH2 domains are required for both heterodimerization
CC with other Bcl2 family members and for repression of cell death.
CC -1- PTM: Proteolytically cleaved by caspases during apoptosis. The
CC cleaved protein, lacking the BH4 domain, has pro-apoptotic
CC activity.
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 1 (BH1).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 2 (BH2).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 3 (BH3).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 4 (BH4).
CC -1- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z23116; CAAB0662.1; -;
CC EMBL: Z23115; CAAB0661.1; -;
CC EMBL: U72398; AAB17354.1; -;
CC PDB: 1BXL; 29-OCT-97.
CC PDB: 1LXL; 21-APR-97.
CC PDB: 1MAZ; 21-APR-97.
CC MIM: 600039; -;
CC InterPro: IPR002475; BCL2_family.
CC InterPro: IPR003093; BCL2.
CC InterPro: IPR000712; BCL2.
CC Pfam: PF00452; BCL2; 1.
CC Pfam: PF02180; BCL1; 1.
CC SMART: SM00337; BCL; 1.
CC SMART: SM00265; BH4; 1.
CC PROSITE: PS50062; BCL2_FAMILY; 1.
CC PROSITE: PS01080; BH1; 1.
CC PROSITE: PS01258; BH2; 1.
CC PROSITE: PS01259; BH3; 1.
CC PROSITE: PS01260; BH4_1; 1.
CC PROSITE: PS50063; BH4_2; 1.
CC Apoptosis: Mitochondrion; Alternative splicing; Transmembrane;
CC 3D-structure.
CC DOMAIN 4 24
CC FT DOMAIN 86 100 BH4.
CC FT DOMAIN 129 148 BH3.
CC FT DOMAIN 180 195 BH1.
CC FT TRANSMEM 210 226 BH2.
CC FT SITE 61 61 POTENTIAL.
CC FT VARSPLIC 126 188 CLEAVAGE BY CASPASE-1.
CC FT VARSPLIC 189 233 MISSING (IN ISOFORM BCL-X(S)).
CC FT VARSPLIC 189 233 DFEVELIGNNAASRQGFNRWFLTGMTAGVAVLLGSL
CC FT VARSPLIC 189 233 FSRK -> VRTKPLCFSLRSGSPALLYLFLLCVIL
CC FT VARSPLIC 189 233 VGVDS (IN ISOFORM BCL-X(BETA)).
CC FT MUTAGEN 61 61 D->A: NO CLEAVAGE BY CASPASE-1 NOR BY
CC FT MUTAGEN 131 133 CASPASE-3.
CC FT MUTAGEN 135 137 FRD->VRA: NO HETERODIMERIZATION WITH BAX.
CC FT MUTAGEN 138 140 VNM->ALL: LOSS OF ANTI-APOPTOTIC
CC FT MUTAGEN 138 140 ACTIVITY.
CC FT MUTAGEN 138 140 GRI->ELN: LOSS OF ANTI-APOPTOTIC
CC FT MUTAGEN 138 140 ACTIVITY.
CC FT MUTAGEN 138 138 G->A: NO HETERODIMERIZATION WITH BAX.
CC FT MUTAGEN 148 148 G->E: NO HETERODIMERIZATION WITH BAX.
CC FT MUTAGEN 156 156 D->A: NO EFFECT ON CASPASE-1 CLEAVAGE.
CC FT MUTAGEN 176 176 D->A: NO EFFECT ON CASPASE-1 CLEAVAGE.
CC FT MUTAGEN 188 189 WD->GA: REDUCES ANTI-APOPTOTIC ACTIVITY
CC FT MUTAGEN 189 189 BY ABOUT HALF.
CC FT MUTAGEN 189 189 D->A: NO EFFECT ON CASPASE-1 CLEAVAGE.
CC FT CONFLICT 70 70 G -> A (IN REF. 1; CAAB0661).

SO SEQUENCE 233 AA; 26049 MW; E09D3CDD851AE9BE CRC64;
Query Match 6.8%; Score 117; DB 1; Length 233;
Best Local Similarity 20.8%; Pred. No. 0.047;
Matches 54; Conservative 35; Mismatches 80; Indels 90; Gaps 12;
QY 77 SQSSSKAI-----NLGKKRSWKAFEGVYER-----EDSQSTPAKVASQORTLEY 122
DB 2 SQSNELVDFSLYSLQSGYMSQSDYENETEAEGESEMETSAING----- 53
QY 123 QDSHSQMSRCLSNVEQCLEHEAVDPKVISIANRVAEIVYSWPPATQAGFSKELFV 182
DB 54 ---NPSW-----HLADSPAV----- 65
QY 183 TEGLSFQLOGHVPVASSSKKDEEQL-LAKIVELLYSGDQLF---KKLKKALMGHFQ 238
DB 66 -NGAT-----GH-----SSSLDAREVIPMAVKKQALREAGDEFLURYRRARSDLTSOLHT 115
QY 239 DGLSYVEKTIIDQVILM-GVDPGESESEVKAQGFKAALVIDVTAKLTIDNHPMNRVLGFG 297
DB 116 PGTAFQSFQVYVNELEFRQGVN--WGRIVAFESFGALCVESVDKEMGY---LVSRIAMM 170
QY 298 TKYKENFSPIQOHGWE 316
DB 171 ATYLDHLEPWIQENGWMD 189
RESULT 5
ID TDT_MOUSE STANDARD; PRT: 530 AA.
AC P09838; Q99PD0; Q99PD1;
DT 01-MAR-1989 (Rel. 10; Created)
DT 01-MAR-2002 (Rel. 41; Last sequence update)
DT 01-MAR-2002 (Rel. 41; Last annotation update)
DE DNA nucleotidyltransferase (EC 2.7.7.31) (Terminal addition enzyme)
DE (Terminal deoxynucleotidyltransferase) (TDT) (Terminal transferase).
GN DNTT OR TDT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86286588; Pubmed=3755527;
RA Koizumi O., Yokota T., Kageyama T., Hirose T., Yoshida S., Arai K.;
RT "Isolation and characterization of bovine and mouse terminal
RT deoxynucleotidyltransferase cDNAs expressible in mammalian cells.";
RL Nucleic Acids Res. 14:5777-5792(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIR/C; TISSUE=Thymus;
RX MEDLINE=93215079; Pubmed=8464703;
RA Doyen N., Fanton D'Andon M., Bentolila L.A., Nguyen T.O., Rougeon F.;
RT "Differential splicing in mouse thymus generates two forms of
RT terminal deoxynucleotidyl transferase.";
RL Nucleic Acids Res. 21:1187-1191(1993).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
RC STRAIN=C57BL/6; TISSUE=Thymus;
RX PubMed=11136823;
RA Benedict C.L., Gilliland S., Kearney J.F.;
RT "The long isoform of terminal deoxynucleotidyl transferase (Tdtl)
RT enters the nucleus and, rather than catalyzing N addition, modulates
RT the catalytic activity of the short isoform.";
RL J. Exp. Med. 193:89-99(2001).
RN [4]
RP PRELIMINARY CHARACTERIZATION OF ALTERNATIVE FORMS.
RX MEDLINE=96016194; Pubmed=755063;
RA Bentolila L.A., Fanton D'Andon M., Nguyen T.O., Martinez O.,
RA Rougeon F., Doyen N.;
RT "The two isoforms of mouse terminal deoxynucleotidyl transferase
RT differ in both the ability to add N regions and subcellular

```

RT Localization.
RL EMBL J. 14:4221-4229(1995).
CC -1- FUNCTION: TEMPLATE-INDEPENDENT DNA POLYMERASE WHICH CATALYZES THE
CC RANDOM ADDITION OF DEOXYNUCLEOSIDE 5'-TRIPHOSPHATE TO THE 3' END OF
CC A DNA INITIATOR. ONE OF THE IN-VIVO FUNCTION OF THIS ENZYME IS THE
CC ADDITION OF NUCLEOTIDES AT THE JUNCTION (N REGION) OF REARRANGED
CC HEAVY CHAIN AND T CELL RECEPTOR GENE SEGMENTS DURING THE
CC MATURATION OF B AND T CELLS. The TdtL isoform seems to serve to
CC modulate the function of Tdts.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate +
CC (deoxynucleotide)(M) - N diphosphate + (deoxynucleotide)(M+N).
CC -1- COFACTOR: REDUCES MAGNESIUM.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: Tdt-L(arge)/TdtL (shown here)
CC and Tdt-S(mall)/Tdts, are produced by alternative splicing. The
CC Tdt-S form is the major form.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-X FAMILY.
CC -1- SIMILARITY: CONTAINS 1 BRC1 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04123; CAA27735.1; -
DR EMBL: X68670; CAA48634.1; -
DR EMBL: AF316014; AAK07884.1; -
DR EMBL: AF316015; AAK07885.1; -
DR PIR: B23595; B23595.
DR HSSP: P06766; 1BPB.
DR MGD: MGI:98659; DntL.
DR InterPro: IPR001357; BRC1.
DR InterPro: IPR002054; DNA_PolX.
DR Pfam: PF00533; BRC1; 1.
DR Pfam: PF00966; DNA_PolymeraseX; 1.
DR PRINTS: PR00869; DNAPOLX.
DR SMART: SM00292; BRC1; 1.
DR SMART: SM00483; POLX; 1.
DR PROSITE: PS50172; BRC1; 1.
DR PROSITE: PS50522; DNA_PolymeraseX; 1.
DR TRANSIT: Nucleotidyltransferase; Terminal addition; Magnesium;
KW Nuclear protein; Alternative splicing.
FT DOMAIN 27 118 MISSING (IN ISOFORM TDT-S).
FT VARSPIC 483 502 T -> M (IN REF. 2).
FT CONFLICT 36 36 L -> F (IN REF. 2).
FT CONFLICT 99 99 L -> F (IN REF. 2).
FT CONFLICT 193 193 R -> G (IN REF. 1).
FT CONFLICT 287 287 Q -> K (IN REF. 1).
FT CONFLICT 309 309 E -> Q (IN REF. 1).
FT CONFLICT 367 367 D -> H (IN REF. 1).
FT CONFLICT 441 444 DRRA -> ECAC (IN REF. 1).
FT CONFLICT 443 445 RAE -> AS (IN REF. 2).
FT SEQUENCE 530 AA; 60331 MW; 66B109DCFC9C8107 CRC64;
SQ

```

Query Match 6.7%; Score 115; DB 1; Length 530;

Best Local Similarity 24.0%; Pred. No. 0.19; Indels 60; Gaps 16;

Matches 71; Conservative 49; Mismatches 116;

```

QY 8 DLEETPLDDDLTIEFKILAYTRHHVFKSTPALFSPKLTSTRSLASRGLCNCSANESM 67
DB 170 DALDILAENDELKENSCLAFMRASSVLKSLPPI-----TSKKDTEGI-PCLDGDKX 222
QY 68 TEVSWPCRNSSSE-KAINLGKKKSSWKAFA--FGVEKEDSQSTPAVSAOQRTL-EY 122
DB 223 STIEGIIIEGSESEAKAVLNDERKSKFLFVSFGVGLK-----TAERWFRMGFTLSKI 277
QY 123 QDSHSGQMSRC-----LSNVEQCLEHENDPKVVISIANVAELIYSG-PPQATQAG 174
DB 278 QSDKSLRFTQOMKAGFLYEDLVSCVNR---PEAEAVSMLVKEAVVTFPLDALTMTGG 333

```

```

QY 175 FKSKEIFVTEGLSFOLOGH-----VPVASSKKDEEOLIAKIVELLYKSG-----DQLER 225
DB 334 FRRC-----KMGHDVDFLTSPATEDEEQQLHKHTDFMKQGLLYLCILES 383
QY 226 KLRKDKA-----LMGHFQDGLSYSEVFTTQDVLMGVDPRESEVKAQGFALY 275
DB 384 TFEFKOPSRKVDALDHFOK--CFILK--LDHGRVHSEKSGQEGK--GMAIRV 433

RESULT 6
BCLX_PIG ID BCLX_PIG STANDARD; PRT; 233 AA.
AC 077737;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Apoptosis regulator Bcl-x.
GN BCL2L1 OR BCL2L OR BCLX.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Bartling B., Hoffmann J., Holtz J., Schulz R., Heusch G., Darmer D.;
RT "Expression of apoptosis-associated genes in hibernating and stunned
RT myocardium of pig";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Potent inhibitor of cell death. Inhibits activation of
CC caspases (By similarity). Appears to regulate cell death by
CC blocking the voltage-dependent anion channel (VDAC) by binding
CC to it and preventing the release of the caspase activator,
CC cytochrome c, from the mitochondrial membrane.
CC -1- SUBUNIT: Bcl-x(L) forms heterodimers with BAX, BAK and Bcl-2 (By
CC similarity). Heterodimerization with BAX does not seem to be
CC required for anti-apoptotic activity (By similarity).
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MEMBRANES AND PERINUCLEAR
CC ENVELOPE (By similarity).
CC -1- DOMAIN: The BH4 domain is required for both heterodimerization
CC with other Bcl2 family members and for repression of cell death.
CC with other Bcl2 family members and for repression of cell death.
CC -1- PTM: Proteolytically cleaved by caspases during apoptosis (By
CC similarity). The cleaved protein, lacking the BH4 domain, has pro-
CC apoptotic activity (By similarity).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 1 (BH1).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 2 (BH2).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 3 (BH3).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 4 (BH4).
CC -1- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ001203; CAA04597.1; -
DR HSSP: Q07817; IMAZ.
DR InterPro: IPR002475; BCL2_family.
DR InterPro: IPR003093; BH4.
DR InterPro: IPR000712; BCL_2.
DR Pfam: PF00452; Bcl-2; 1.
DR Pfam: PF02180; BH4; 1.
DR SMART: SM00357; BCL; 1.
DR SMART: SM00265; BH4; 1.
DR PROSITE: PS50062; BCL2_FAMILY; 1.
DR PROSITE: PS01080; BCL; 1.
DR PROSITE: PS01258; BH2; 1.
DR PROSITE: PS01259; BH3; 1.
DR PROSITE: PS01260; BH4_1; 1.

```


QY 261 KLTADNHPNVLGFGTKYKLNFSPIQHGWEKIL 319
ID : : : : :
Db 128 KVEL-----IRITMGWTLDFRLRLVWIDOGGMEGL 162

RESULT 12
ARL1_XENLA STANDARD; PRT; 204 AA.
AC 091828;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Apoptosis regulator R11 (XR11).
OC Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Tissue-Head:
RA MEDLINE=95331613; PubMed=7607538;
RA Cruz-Reyes J., Tata J.R.,
RA "Cloning, characterization and expression of two Xenopus bcl-2-like
RA cell-survival genes".
RL Gene 158:171-179(1995).
CC -1- FUNCTION: CONFERS STRONG PROTECTION AGAINST CELL DEATH.
CC -1- FUNCTION: LOCATIONS: Membrane-bound (Potential).
CC -1- DEVELOPMENTAL STAGE: DEVELOPMENTAL REGULATION ONLY OCCURS IN THE
CC BRAIN OF MID-METAMORPHOSIC TO POST-METAMORPHOSIC TADPOLES AND
CC ADULTS, WHERE AN INCREASE OF SEVERAL FOLD HAS BEEN OBSERVED.
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 1 (BH1).
CC -1- SIMILARITY: CONTAINS 1 BCL-2 HOMOLOG DOMAIN 2 (BH2).
CC -1- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X82461; CAA57844.1; -
DR HSSP: Q07817; 1MAZ.
DR InterPro: IPR002475; BCL2_family.
DR InterPro: IPR003093; BH4.
DR InterPro: IPR000712; BCL_2.
DR Pfam: PF00452; Bcl-2; 1.
DR Pfam: PF02180; BH4; 1.
DR SMART: SM00337; BCL; 1.
DR SMART: SM00265; BH4; 1.
DR PROSITE: PS01080; BH1; 1.
DR PROSITE: PS01258; BH2; 1.
DR PROSITE: PS50062; BCL2_FAMILY; 1.
DR Apoptosis; Transmembrane.
DR DOMAIN 101 120 BH1.
DR DOMAIN 152 167 BH2.
DR TRANSMEM 181 198 POTENTIAL.
FT TRANSMEM 181 198
SQ SEQUENCE 204 AA; 23379 MW; 3BFC6BEDDAACA03 CRC64;

Query Match 5.4%; Score 93; DB 1; Length 204;
Best Local Similarity 20.6%; Pred. No. 2,7;
Matches 43; Conservative 34; Mismatches 58; Indels 74; Gaps 9;

QY 131 SRCLSNVEOCLHEAVDPKVISIANRYAEIVYSPPPQATGAGGFKSEIVTEGLISFOL 190
ID : : : : :
Db 14 SKRLSQNEACRK-----FSNNPNPMPEYLMEPTSERP-----EGAT--- 50

QY 191 QGHVPAVASSSKKEEEOILAKIYVELLKYSQDLERKTLKKDAIMGHFDGLSYSEFKTTI 250
ID : : : : :
Db 191 QGHVPAVASSSKKEEEOILAKIYVELLKYSQDLERKTLKKDAIMGHFDGLSYSEFKTTI 250

Db 51 QGIV-----EEVLQALLATE-----EELRYQRA-----FSDLT 81

QY 251 DQVLAVGDPRESEVEYKAG-----FKALVIVDYAK-LTAIDNHPNR 292
ID : : : : :
Db 82 SQLIHTIDDTAQSGSQQVVGELFRODGTNMGRIYAFPSFGALCVESANKENTDL-----LPR 137

QY 293 VLGFQTKYKLNFSPIQHGWEKIL 321
ID : : : : :
Db 138 IVQMWVNLLEHTLPQPMQENGMEAFVGL 166

RESULT 13
PMS2_MOUSE STANDARD; PRT; 859 AA.
AC P54279;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE PMS1 protein homolog 2 (DNA mismatch repair protein PMS2).
GN PMS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=95354212; PubMed=7628019;
RA Baker S.M., Bronner C.E., Zhang L., Plug A., Robatzek M.,
RA Warren G., Elliot E.A., Yu J., Ashley T., Antheim N.,
RA Flavell R.A., Liskay R.M.,
RA "Male mice defective in the DNA mismatch repair gene PMS2 exhibit
RA abnormal chromosome synapsis in meiosis".
RL Cell 82:309-319(1995).
CC -1- FUNCTION: PROBABLY INVOLVED IN THE REPAIR OF MISMATCHES IN DNA.
CC -1- SUBUNIT: HETERODIMER OF PMS2 AND MLH1.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEXB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U28724; AAA87031.1; -
DR HSSP: P23367; 1BKN.
DR MGD: MGI:104288; Pms2.
DR InterPro: IPR002099; DNA_mis_repair.
DR InterPro: IPR003594; HATPase_C.
DR Pfam: PF01119; DNA_mis_repair; 1.
DR Pfam: PF02518; HATPase_C; 1.
DR PROSITE: PS00058; DNA_MISMATCH_REPAIR_1; 1.
DR DNA_repair.
KW DNA_repair.
SQ SEQUENCE 859 AA; 95225 MW; 263B5A6BBB2ACA9 CRC64;

Query Match 5.4%; Score 93; DB 1; Length 859;
Best Local Similarity 22.2%; Pred. No. 17;
Matches 61; Conservative 46; Mismatches 124; Indels 44; Gaps 10;

QY 53 LSQRLG-GNCSANESWTEVSWPCRNSSSEKAINIGKKSSWKAFFGVYKEKDSQTPAK 111
ID : : : : :
Db 463 ISYGLRSQDKIVSPDSDPCMDREKIEKDSGLSTAGSEEFSTPEVASSFSDDYN 522

QY 112 VSAGQGRLEBQDSHSQWMSR-----LSNVEQL--EHEANDPVIISIANRYAEIVY 163
ID : : : : :
Db 523 VSS-----LEDRPSQITNGCIDCRPPGTGQSLKPPDHGYQCKALPLARL----- 568

QY 164 WPPQATGAGGFKSEIVTEGLSFOLQGHVPAVASSSKKEEEOILAKIYVELLKYSQDL 223
ID : : : : :
Db 569 -----SPTNAKFKFTFEERSNVNISQRLRG--PQSTSAALVDAIKMKRIVLLEFSLSSL 622

QY 224 ERRLKDKALMGHFDGLSYSVFKTTIDVLMGVDPGESEVKAQGFRAALY-IDVTAKL 282
DB 623 AKRMKQLOHLKAKQNKHLSTYKFRFA---KICPGEMQAEDELRKELSKSMFAMELIGOF 679
QY 283 TAIDNHPMNRVILGFGTKYKLENFSPWLOQHGWEK 317
DB 680 N-----LGFIVTKLEKEDLF-LVDQHADEK 703
RESULT 14
GGA3_HUMAN STANDARD; PRT: 723 AA.
ID GGA3_HUMAN Q9N552; Q9JUY3; Q15017;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DEF ADP-ribosylation factor binding protein GGA3 (Golgi-localized, gamma
ear-containing, Arf-binding protein 3).
GG3 OR KIA0154.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS).
RC TISSUE=Heart;
RX MEDLINE=20211638; PubMed-10747089;
RA Dell'Angelica E.C., Puertollano R., Mullins C., Aguilar R.C.,
Vargas J.D., Hartwell L.M., Bonifacio J.S.;
"Ggas" a family of Adp-ribosylation factor-binding proteins related to
adaptors and associated with the Golgi complex.";
J. Cell Biol. 149:81-94(2000).
[2]
SEQUENCE FROM N.A. (SHORT ISOFORM).
RX MEDLINE=20214818; PubMed-10749927;
RA Boman A.L., Zhang C.-J., Zhu X., Kahn R.A.;
"A family of Adp-ribosylation factor effectors that can alter
transport through the trans-Golgi."
Mol. Biol. Cell 11:1241-1255(2000).
[3]
SEQUENCE FROM N.A. (SHORT ISOFORM).
RP TISSUE=Bone marrow;
RX MEDLINE=96127530; PubMed-8590280;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
"Prediction of the coding sequences of unidentified human genes. IV.
The coding sequences of 40 new genes (K1A0121-K1A0160) deduced by
analysis of cDNA clones from human cell line KG-1."
DNA Res. 2:167-174(1995).
-1- FUNCTION: MAY PLAY A ROLE IN THE REGULATION OF MEMBRANE TRAFFIC
THROUGH THE TRANS-GOLGI NETWORK.
-1- SUBUNIT: BINDS TO ARF1.
-1- SUBCELLULAR LOCATION: TRANS-GOLGI NETWORK.
-1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
-1- TISSUE SPECIFICITY: UBIDITUOUSLY EXPRESSED.
-1- SIMILARITY: CONTAINS 1 GAMMA-ADAPTIN C-TERMINAL DOMAIN.
-1- SIMILARITY: CONTAINS 1 VHS DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb-sdb.ch).
CC
DR EMBL: AF219138; AAF42848.1; -
DR EMBL: AF190864; AAF05709.1; -
DR EMBL: AF219139; AAF42849.1; -
DR EMBL: D63876; BAA09926.1; ALT_INIT.
DR InterPro: IPR001121; G_adapt_C.
DR InterPro: IPR002014; HRS.
DR Pfam: PF00790; VHS; 1.

DR Prodom; PD003686; HRS; 1.
DR Prodom; PD021457; G_adapt_C; 1.
DR SMART: SMO0288; VHS; 1.
DR PROSITE, PS50179; VHS; 1.
KW Protein transport; Golgi stack; Coiled coil; Alternative splicing.
FT DOMAIN 1 313 BINDS TO ARF1 (IN LONG ISOFORM).
FT DOMAIN 16 146 VHS.
FT DOMAIN 188 232 COILED COIL (POTENTIAL).
FT DOMAIN 357 360 POLY-PRO.
FT DOMAIN 453 457 POLY-SER.
FT DOMAIN 598 709 POLY-VAL.
FT VARSPIC 68 100 GAMMA-ADAPTIN C-TERMINAL.
SQ SEQUENCE 723 AA; 78314 MW; 4F80D6032239168C CRC64;
Query Match 5.4%; Score 92.5; DB 1; Length 723;
Best Local Similarity 18.4%; Pred. No. 15;
Matches 65; Conservative 49; Mismatches 85; Indels 155; Gaps 17;
QY 65 ESWTEVSPPCRRNSQSEKALNGLKSSKKAFFGVYEKEDSOSTPAKVSAGQRTLEYQD 124
DB 10 ESWL-----NKAATPSNRQEDWEYIIIGFDQINKLEGPQIAV---RLAKKI 54
QY 125 SHSQQWS--RCLSNVEQCLEH-----EAVDPKYI-----SIANRV 157
DB 55 QSPQEWELQALTYLEACMKNGRRHNEVGKRFLELIKVSPKILGDRVSEKATKY 114
QY 158 AEIYYSW-----PPQATQAGGFSKE 179
DB 115 IELYSWTMALPEEAKIKDAYHMLKRGQIVQSDPPIPVDRTLIPSPPR-----PKNP 167
QY 180 IEVTEGLSFQI-----QGHVYVASSKDE-----EQ-----QI 208
DB 168 VFDEEKSKLAKLIKSKNPDLDQENAKLIKSNVKKDEARIQVTKRLHTLEEVNVRNL 227
QY 209 LAKIV-----ELKYSQDQLERKIK-----KDKALMGHFO--DGLS 242
DB 228 LSEMLIHYSDQESSDGDRELMKEIPDQCKNKRITLKLASETEDNDNSLDIILQASNDLS 287
QY 243 YSV--FKTTID-QVLMG-----VDPGESEVKAQGFRAALY-IDVTAKLTAI 285
DB 288 RVINSYKTIIEGVINGEVATLTLPDSEGNQCSNGGTILDLAELDTNTLSLSV 341
RESULT 15
RPN2_YEAST STANDARD; PRT: 945 AA.
ID RPN2_YEAST P32565;
AC P32565;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 26S proteasome regulatory subunit RPN2.
GN RPN2 OR SEN3 OR Y11075C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CC NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=96026012; PubMed=7565784;
RA Demarini D.J., Papa F.R., Swaminathan S., Ursic D., Rasmussen T.P.,
Culbertson M.R., Hochstrasser M.;
"The yeast SEN3 gene encodes a regulatory subunit of the 26S
RT proteasome complex required for ubiquitin-dependent protein
RT degradation in vivo."
RL Mol. Cell. Biol. 15:6311-6321(1995).
[2]
SEQUENCE FROM N.A.
RP STRAIN=S288C / AB972;
RC Bartell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagers K., Jones M.,

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 16:32:27 ; Search time 31.42 seconds
(without alignments)
1000.039 Million cell updates/sec

Title: US-09-771-961-2

Perfect score: 1709

Sequence: 1 MCSYSGCDLEIPIIDDDDLN.....WIOQHGWEKILISHEVD 327

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 283138

Database : PIR_71:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Result No.	Score	Query Match	Length	ID	Description
1	120	7.0	233	2 S51761	BCL-X protein - ra
2	119.5	7.0	190	2 A47537	apoptosis regula
3	119	7.0	233	2 I49056	bcl-x long - mouse
4	118	7.0	233	2 I67431	bcl-x long - rat
5	118	6.9	214	2 I49057	bcl-x transmembran
6	117	6.8	233	2 B47537	apoptosis regula
7	115	6.7	227	2 JEO033	apoptosis regula
8	115	6.7	562	2 T49904	hypothetical prote
9	113.5	6.6	1647	2 T49412	hypothetical prote
10	113	6.6	529	2 B23595	DNA nucleotidylexo
11	105.5	6.2	901	2 H64101	preprotein transio
12	104	6.1	414	2 A29835	Sali protein - Esc
13	100	5.9	2052	2 C97038	phage-related prot
14	99	5.8	757	2 C70034	conserved hypotet
15	97.5	5.7	567	2 G95233	Cof family protein
16	97	5.7	492	2 G70403	lysine--trna ligas
17	95.5	5.6	1055	2 AD2499	hypothetical prote
18	95	5.6	699	2 AC0504	probable chitinase
19	94.5	5.5	192	2 D47538	bcl-2-associated p
20	94	5.5	561	2 T22917	probable ATP-depen
21	94	5.5	1123	2 A72311	conserved hypotet
22	93	5.4	450	2 T37702	hypothetical prote
23	93	5.4	670	2 F84540	hypothetical prote
24	92.5	5.4	945	2 S48369	26S proteasome reg
25	92	5.4	333	2 T04656	hypothetical prote
26	92	5.4	880	2 E69680	DNA polymerase I p
27	92	5.4	1083	2 T00790	ubiquitin-specific
28	92	5.4	1384	2 T52301	GYMOS/PICKLE prot
29	91.5	5.4	547	1 A23595	DNA nucleotidylexo

30	91	5.3	875	2 S70115
31	91	5.3	1660	2 A84647
32	90.5	5.3	472	2 A99098
33	90.5	5.3	878	2 A83748
34	90	5.3	433	2 S63143
35	89.5	5.2	218	2 B47538
36	89.5	5.2	1029	2 F86210
37	89.5	5.2	2319	2 A47004
38	89	5.2	926	2 E64171
39	89	5.2	1722	1 T78879
40	89	5.2	2253	2 T30336
41	88.5	5.2	445	2 B40590
42	88.5	5.2	578	2 D90564
43	88.5	5.2	1333	2 E84601
44	88.5	5.2	1359	2 B84645
45	88.5	5.2	1447	2 T00530

ALIGNMENTS

RESULT 1

S51761

BCL-X protein - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 02-Mar-2001

C:Accession: S51761; S51762

R:Michaelidis, T.M.

Submitted to the EMBL Data Library, November 1994

A:Reference number: S51761

A:Accession: S51761

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-233 <MTC>

A:Cross-references: EMBL:X82537; NID:9607176; PIDN:CAA57886.1; PID:9607177

A:Experimental source: embryonic; brain

A:Accession: S51762

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-125,189-233 <MT2>

A:Cross-references: EMBL:X82537; NID:9607176; PIDN:CAA57887.1; PID:9607178

A:Experimental source: embryonic; brain

A>Note: smaller form due to splicing

C:Genetics:

A:Introns: 125/3

C:Superfamily: bcl transforming protein

Query Match

Best Local Similarity 7.0%; Score 120; DB 2; Length 233;

Matches 54; Conservative 39; Mismatches 81; Indels 90; Gaps 12;

QY	77	SQSEKAI-----NLGKKSSWKAFFGVYK-----EDSQSPAKYSAQOQTLEY	122
DB	2	SQSQDELVDLSTYKLSQKGYMSQSFSDVENRTPEETEPREPSAING-----	53
QY	123	QDSHQWQRCLSNVQCLEHFAVDPKVSIANRAVEIYSWPPQATQAGFKSKIEFV	182
DB	54	-----NPSW-----HLADSPAV-----	65
QY	183	TEGLSFQLOGHVPYASSSKKDEEIOI-LAKIVELTKYSGDLE---RLKDKRALMGFQ	238
DB	66	-NGAT-----GH-----SSLDAREVIPAAMAVQALNEAGDEFELRRARRASDLSQHLIT	115
QY	239	DGLSYVFKTTTDOYLM-GVDPGRGESEVKAQGFKAALVTDYAKLPAIDNHPNRYLGFQ	297
DB	116	PGLAYQSEFQOVNNELEFQGVN--WGRIVAFESFGALCVESVDKEMQV--LVSRIASWM	170
QY	298	TKYLKENFSPMIOHGWKEKILGI	321
DB	171	ATYLDHLEPWIQENGWDTFVDL	194

RESULT 2
A47537
apoptosis regulator bcl-x - chicken
C:Species: gallus gallus (chicken)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Jul-1999
C:Accession: A47537
R:Boise, L.H.; Gonzalez-Garcia, M.; Postema, C.E.; Ding, L.; Lindsten, T.; Turka, L.A.;
Cell 74, 597-608, 1993
A:Title: bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptosis
A:Reference number: A47537; MUID:93364977
A:Accession: A47537
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <BOI>
A:Cross-references: GB:J23110; GB:I20120; NID:g510898; PIDN:CAA80657.1; PID:g510899
C:Superfamily: bcl transforming protein

	Query Match	7.0%	Score 119.5	DB 2	length 190	
	Best Local Similarity	22.4%	Pred. No. 0.041			
	Matches	48	Conservative	38	Mismatches	65
					Indels	43
					Gaps	8
Qy	113	SAAGGRTLEYDSDSHSOQSRCLSNVEQCLIEH---	AVDPKVTISIANRAEIVYSWPPQA	169		
Db	3	SSNRRLVIDFVSYKLSQRGHCSWSEEDENRDTAAEAKMSYVLSGP---	SWMP--	56		
Qy	170	TQAGGFKSKSEIVTEGLSPQLQGHV---	PVSSSSKKDEEDILAKIYELLYKSGPQLE-	224		
Db	57	-----AGHYVNCATVYRSSLSEHVEIYRASGVKALDADGDEFLL	95			
Qy	225	--RLKKDKALMGHFQDGLSYVSFEKTTIDQVLM-	GVDPGRGESEVKAQAGKAAALVIDYTK	281		
Db	96	RYRRAFSDLTSLQHTTPGATAYOSFEQVAVELFHDGYN-	WGRIVAFEFSGALCVESYDK	153		
Qy	282	LTAIDNHPRNRYLGGCTKYLKENESPMIOQHCGW	315			
Db	154	EMRV--LVGRIVSWMTLYLTDLHLDPMVIOENCGW	184			

RESULT 3
149056 bcl-x long - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
C:Accession: 149056, S52866
R:Fang, W.; Rivard, J.J.; Mueller, D.L.; Behrens, T.W.
J. Immunol. 153, 4388-4398, 1994
A:Title: Cloning and molecular characterization of mouse bcl-x in B and T lymphocytes.
A:Reference number: 149055, MUID:95052604
A:Accession: 149056
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-233 <RES>
A:Cross-References: EMBL:U010101; NID:g506647; PIDN:AAA82173.1; PID:g506648
R:Kamesaki, H.; Michaud, G.Y.; Takatsu, K.; Okuma, M.
Submitted to the EMBL Data Library, November 1994
A:Description: IL-5 inhibits anti-IgM-induced apoptosis in an immature B cell line thro
A:Reference number: S52866
A:Accession: S52866
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-233 <KAM>
A:Cross-References: EMBL:X83574; NID:g695622; PIDN:CAA58557.1; PID:g695623
C:Superfamily: Bcl transforming protein

```

Query Match 119; DB 2; Length 233;
Best Local Similarity 21.1%; Pzed No. 0.05;
Matches 54; Conservative 37; Mismatches 91; Indels 74; Gaps 11

OY 77 SOSSEKAT-----NLGKKSSWKAFEGVEEKEDSOSTPAKYSAGQRLFLQDSHSQOM 130
    111: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2 SOSNELVYDFLSYKLSQKGYMSQSDVEENR---TAPETERERETPSPAINPSPM 57

```

```
QY      131 SRCLSNEQCCEHAEVADPKVISIANRVAELIYSMPPEQTAGGFKSEIFVEGSLFOL 190
          | | | | |
Db      58 -----HLADSPAV-----NGAR---69

QY      191 QGHVPVASSSKKDEEOI-LAKIVELLKYSGDOL--RLKKDKALMGHFODGLSYSVF 246
          | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db      70 -GH-SSSLDAREVIYPMAAAYQALREAGDEFELRYRRASFDSLTSQLHTPGTAQSOF 123
```

RESULT 4
167431
BCL-X-Long - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 16-Jul-1999
C:Accession: 167431
R:Tilly, J.L.; Tilly, K.I.; Kenton, M.L.; Johnson, A.L.
Endocrinology 136, 232-241, 1995
A:Title: Expression of members of the bcl-2 gene family in the immature rat ovary: equi-
constitutive bcl-2 and bcl-X-long messenger ribonucleic acid levels.
A:Reference number: 153295; MID:95129487
A:Accession: 167431
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-233 <RES>
A:Cross-references: EMBL:U34963; NID:g1004376; PIDN:AAA77686.1; PID:g1004377
A:Superfamily: bcl transforming protein

	Query Match	7.0%; Score 119; DB 2; Length 233;
	Best Local Similarity	20.2%, Pred. No. 0.059;
Matches	53; Conservative	38; Mismatches 64; Indels 88; Gaps 11.
QY	77 SSSSEKAI-----NLGKKSKSMKAFFGVGEK-----EESOSNPAKVSAOQGRTLEY	122
Dd	2 SOSNELVVDLSTYLKSOKGYSWSOFSDEENKTAPTEPTEPERETPSAING-----	53
QY	123 QDSHSQQSRCLSNVEQCLEHEAVDPKYISIANRYAEIVYSWPPQAOTAGGFCKSEIPV	182
Dd	54 -----NPSPM-----HLADSPAIV-----	65
QY	183 TEGSLFQLGHVPVAASSKRDEEQILAKIYEVLTKSGDOL--RKLLKKDALMGHFOOD	239
Dd	66 -NGAT----GH---SSSIDAREVLPMAAVKOALREAGDEFELRYPRAASDLTSQLHTRP	116
QY	240 GLSYIVEFTITDOULM-GVDFRGEESEVAAOGKALKALVIDVTAKLAIDNHRRMYLGFGCT	298
Dd	117 GTVVVSFEQVYNNEIRPDGGVN-WGRIVASSSFPGLACVESVDKEMOV--LVLSRIASMSMA	171
QY	299 KYIKENESPWIOOHGGMWEKITIGI	321
Dd	172 TYLANDHPEDWGMDTFEVDL	194

RESULT 5
I49057
bcl-x transmembrane deleted - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revission 02-Jul-1996 #text_change 16-Jul-1999
C:Accession: I49057
R:Fang, W.; Rivard, J.Y.; Mueller, D.L.; Behrens, T.W.
J. Immunol. 153, 4388-4398, 1994
A:Title: Cloning and molecular characterization of mouse bcl-x in B and T lymphocytes
A:Reference number: I49055, MIMD:95052604
A:Accession: I49057

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-562 <BEV>
A:Cross-references: EMBL:AL351013; GSPDB:GN00063; ATSP:T24H18.70
C:Experimental source: cultivar Columbia; BAC clone T24H18
C:Genetics:
A:Gene: ATSP:T24H18.70
A:Map position: 5
A:Introns: 73/3

Query Match 6.7%; Score 115; DB 2; Length 562;
Best Local Similarity 21.7%; Pred. No. 0.4;
Matches 80; Conservative 43; Mismatches 136; Indels 110; Gaps 16;

QY 31 TRHHFKSPALFSPKL-----LRTSL-----ORGLGCSANESWTEVS 71
DB 128 SHHHYTMDSALRSPFLHFTTGRSGSVDFRSVSCNDYKKGKGFDRKSLKSNLVV- 166
QY 72 WPCRNQS--SEKAINLG-----KKSSWKAFF--GVVEKEDSOSTPAKVSA 114
DB 187 -PLTDSHSAVVSQPPNRGGRVMSLFPKTKKOKNSIFNPSITEKSEVEVLKDSG 245
QY 115 OGQRLEUQ-----DSHSQMSRCLSNVEQLEH---FAVDPKVISIANRVAETIV 162
DB 246 SGVEKLKRLMEANRSRDALTYOSEKSSIGELSEKLQYLESYCDNLKALREATEV- 304
QY 163 SMPPOATGAG--GPKSKEIFVTEGLSFOLGHPVASS-----SKKDEE 206
DB 305 ----SQENSGGSSGKKNSMPVSE--EVWVEGFLQIVSEARLSIKQFLKLVSELEED 358
QY 207 QILAKIVELLKYSQDGLERRLKKDKALMGHFQDGLSYSVFKTITDOVLG----VDP 260
DB 359 STLINININL-LQPHNLSEFSKYSKIQYHLEALISQSVYODFEKCVFOKNGKPLDPE 417
QY 261 GSEVKAQGFKAALVIDYTKLTAIDNHPRNVLGFCFKYLKENS----- 306
DB 418 QDRQANFSSF-----ASLRNLMSNEVLKKGKYYSDFSRFCDEKMSLITTL 465
QY 307 ----PMIQ 311
DB 466 NMTRRMSEQ 474

RESULT 9
T24H18
hypotheical protein BID4.270 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence-revision 02-Jun-2000 #text-change 02-Jun-2000
Accession: T24H18
hulte, U.; Algen, V.; Hobeisel, J.; Brandt, P.; Farmann, B.; Holland, R.; Nyakatura,
Submitted to the Protein Sequence Database, May 2000
A:Reference number: 225022
A:Accession: T24H18
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1647 <SCH>
A:Cross-references: EMBL:AL351928; GSPDB:GN00116; NCSP:BID4.270
C:Experimental source: BAC clone BID4; strain OR74A
C:Genetics:
A:Gene: NCSP:BID4.270
A:Map position: 6

Query Match 6.6%; Score 113.5; DB 2; Length 1647;
Best Local Similarity 23.2%; Pred. No. 2.3;
Matches 71; Conservative 46; Mismatches 102; Indels 87; Gaps 15;

QY 77 SOSSEKAINLG-----KKSSWKAFFGVVEKEDSOSTPAKVSAQO----- 117
DB 422 SPDSAKVELGQDHVGEGRKESVAEAVVGLAEQADREAGELARAEQSAQAGITQDLE 461
QY 118 --RTLEYODSHSQMSRCLSNVEQLEHE-----AVDPKVISIANRVAETIVSWPPP 167

DB 482 QMSAETDDEFARSQTAQAIOKOLEKDEKNDVLESALDAEV---AKEVRSTV-----D 533
QY 168 QATQAG-----GPKSKEIFVTEG-----LSFOLGHPVAVSSSKKDE-----E 205
DB 534 KAAQAGQGPVADSWESADEIVIEEAPLKYVNFPLKPMITTI-SVETEDERVRYFRE 592
QY 206 EQLAKIVELLKYSQDGLERRLKKDKALMGHFQDGLSYSVFKTITDOVLG----GVDP 262
DB 593 EALID--IARLKKDFQIDRN-----LVSASESYMAVGMASKAGLRVROEDGDKAKLF 644
QY 263 SEVKAQGFKAALVIDYTKLTAIDNHPRNVLG---FGFKYLKENSPPMIQHGHEKTL 319
DB 645 TPTKDRIFNVAI-----SSASNQHPKEAIIIGVSGTIV-----WVLKNGDRDHL 691
QY 320 GISHEE 325
DB 692 EDAPHPE 697

RESULT 10
B23595
DNA nucleotidylexotransferase (EC 2.7.7.31) - mouse
N:Alternate names: terminal addition enzyme; terminal deoxyribonucleotidyltransferase
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1988 #sequence-revision 02-Jun-1988 #text-change 18-Jun-1999
Accession: B23595; S30235
R:Kojima, O.; Yokota, T.; Kageyama, T.; Hirose, T.; Yoshida, S.; Arai, K.
Nucleic Acids Res. 14, 5777-5792, 1986
A:Title: Isolation and characterization of bovine and mouse terminal deoxynucleotidyl
A:Reference number: A93633; MUID:86286588
A:Accession: B23595
A:Molecule type: mRNA
A:Residues: 1-529 <KRO>
R:Dojyen, N.; d'Andon, M.F.; Bentolila, L.A.; Nguyen, Q.T.; Rougeon, F.
Nucleic Acids Res. 21, 1187-1191, 1993
A:Title: Differential splicing in mouse thymus generates two forms of terminal deoxyn
A:Reference number: S30235; MUID:93219079
A:Accession: S30235
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-25 'M', 27-98 'F', 100-192 'R', 194-286 'Q', 288-308 'E', 310-366 'D', 368-44
A:Cross-references: EMBL:X68670; NID:9287808; PIDN:CAA48634.1; PID:9287809
C:Superfamily: DNA nucleotidylexotransferase
C:Keywords: alternative splicing; magnesium; nucleotidyltransferase; nucleus

Query Match 6.6%; Score 113; DB 2; Length 529;
Best Local Similarity 24.0%; Pred. No. 0.53;
Matches 71; Conservative 48; Mismatches 117; Indels 60; Gaps 16;

QY 8 DLEETPLDDDDLTMTIEKLLAYTRHHFKSPALFSPKRLTRLSQGLGCSANESW 67
DB 170 DALDILAENDELRENBESCLAFMGASVYLKSLPPT-----TSKMDTGT-PCLDKVK 222
QY 68 TEVSWPCRNQSSE-KAINEGKKSSWKAFF--FGVVEKEDSOSTPAKVSAQOQTL-EX 122
DB 223 STIEGIEEGESESEAKVANDERKSFRLFSVREVGK-----TAEKFRNGFRRLSKI 277
QY 123 QDSHSQMSRCL-----LSNVEQCLEHAADVPKVISIANRVAETIVYSN-PPQATQAG 174
DB 278 QSDKSLRFTKMQKAGFLYEDLVSCVNR---PEAAVSMLEKVAIVTEPLPALVTMTGG 333
QY 175 FKSKEIFVTEGLSFOLGHPVAVSSSKKDEEQLAKIVELLKYSQ-----QDLER 225
DB 334 FRKG-----KMGCHVDLFITSPATEDEEQDLHKYTHRWKQOGLLYCYILES 383
QY 226 KKKDKA-----LMGHFQDGLSYSVFKTITDOVLGMDVDPRESEVKAQGFKAALV 275
DB 384 TEKFKQPSKVALDHFQK--CFLLIK--LDHGVHSEKSSQOGECK--GMAAIRV 433

C:Superfamily: preprotein translocase secA
C:Keywords: ATP; membrane-associated complex; P-loop; protein transport
F:100-107/Region: nucleotide-binding motif A (P-loop) #status atypical
F:203-208/Region: nucleotide-binding motif B
F:207-210/Region: DEAD motif

212

2

5

10

2

Gaps,

318

S 378

5436

489

538 T

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2002, 04:19:16 ; Search time 2402.36 Seconds
(without alignments)
4264.219 Million cell updates/sec

Title: US-09-771-961-3

Perfect score: 759
Sequence: 1 atgtgtagcaccagtggtggtg.....gttgatggtcgtcatttga 759

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Archived: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*
1: em_estbda:*
2: em_estbhm:*
3: em_estln:*
4: em_estlm:*
5: em_estm:*
6: em_estp:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estz:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	637.8	84.0	980	10	BI833102 603090814
2	603.4	79.5	763	10	BG718587 602696672
3	549	66.7	774	10	BG720221 BG720221
4	506.6	62.7	715	10	BI838169 603083677
5	497.2	65.5	672	10	BG717835 602693855
6	418.4	55.1	703	10	BI908814 603066277
7	356.8	47.0	422	12	AQ131601 HS_3049_A
8	321.4	42.3	406	9	AA399486 zt60c07.r
9	294.6	38.8	1535	11	AK016997 Mus muscu
10	294.6	38.8	1541	11	AK008682 Mus muscu
11	294.6	38.8	1578	11	AK018579 Mus muscu
12	294.6	38.8	2222	11	AK016670 Mus muscu
13	276.2	36.4	776	10	BI463753 603203445
14	246.6	32.5	920	10	BI453888 603174765
15	246.4	32.5	652	9	AA536718 vj88a02.r
16	243.6	32.1	680	10	BG245730 602358519
17	242.6	32.0	448	10	BF198484 248427 NA

18	234.4	30.9	697	10	BF584699 602098438
19	227.2	29.9	924	10	BG176106 602337937
20	209	27.5	666	9	BB628327 BB628327
21	203.8	26.9	1061	10	BF139422 601785402
22	201.2	26.5	403	10	BF342342 370081 MA
23	197	26.0	567	9	AM631581 90021 MAR
24	190.6	25.1	922	10	BB615307 BB615307
25	186	24.5	465	10	BF149673 BF149673
26	149	19.6	423	10	BG895661 359228 MA
27	133.6	17.6	437	9	AA061006 m146f07.r
28	130.6	17.2	697	9	AT614194 v177e07.r
29	126.6	16.7	376	10	BF652080 60329866
30	123	16.2	354	9	AM414655 48366 MAR
31	110.2	14.5	984	9	BB013694 BB013694
32	109.2	14.4	376	9	AM484860 62957 MAR
33	104.8	13.8	370	9	AM428115 64957 MAR
34	97.6	12.9	216	10	BF360692 388694 MA
35	93.8	12.4	332	10	T69681 yca4f12.r1
36	89.4	11.8	473	9	AA400686 zu70c11.r
37	88.4	11.6	446	9	BB746476 BB746476
38	87.4	11.5	256	9	AI606692 m157b10.y
39	84.4	11.1	426	9	AM000827 w45d10.x
40	64.6	8.5	195	9	AA107133 m157b10.r
41	59.4	7.8	451	10	BM433532 10M16H02
42	57	7.5	269	9	BB576539 BB576539
43	50.2	6.6	514	12	AO884233 HS_5504_B
44	48.4	6.4	184	9	AA492696 v177e07.r
45	40.2	5.3	221	9	BB589687 BB589687

ALIGNMENTS

RESULT 1
BI833102 980 bp mRNA linear EST 04-OCT-2001
LOCUS 603090814F1 NIH_MGC_120 Homo sapiens CDNA clone IMAGE:5229819 5',
DEFINITION mRNA sequence.
ACCESSION BI833102.1 GI:15944652
VERSION BI833102
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgs.ncl.nih.gov/
1 (bases 1 to 980)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMN at:
http://image.llnl.gov
Plate: LLM11578 row: k column: 04
High quality sequence stop: 803.
Location/Qualifiers
1..980
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5229819"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is

FEATURES

source

Df	625	GTTTATTCCTGGCACACCAACAAGGACCAGCGAGAGGCATTCAAAGTCCAAAGAATT	684
Oy	541	tttgtaactgaagggtcttcttcacagcccaagccaagtgcctgatcctaattctc	600
Df	685	TTTGTATACTGAAGGCTCCTCTCCATTCAGACTCAAGGCCACAGTGCTTAGCTTCAAGATTCT	744
Oy	601	aagaa gots 	
Df	745	AAGAA 749	
RESULT	3		
BG720221		774 bp mRNA linear EST 08-MAY-2001	
LOCUS		602692227FL NIH_MGC_97 Homo sapiens CDNA clone IMAGE:4824760 5'	
DEFINITION		mRNA sequence.	
ACCESSION		BG720221	
VERSION		BG720221.1 GI:13999408	
KEYWORDS		EST.	
SOURCE		human.	
ORGANISM		Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE		1 (bases 1 to 774)	
AUTHORS		NIH-MGC http://mgc.nci.nih.gov/.	
TITLE		National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: Robert Strausberg, Ph.D. Email: cgrabs-femail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGR), Shiraki Toshituki and Piero Carninci (RIKEN) CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: LHAM10736 row: m column: 17 High quality sequence stop: 769. Location/Qualifiers 1..774 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image=4824760" /clone_id="NIH_MGC_97" /lab_host="DH10B" /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgag) : Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to RQF 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NHMH/NHGRI, National Institutes of Health). Note: this is a NIH-MGC Library."	
FEATURES			
source			
BASE COUNT		225 a 185 c 198 g 166 t	
ORIGIN			
Query Match		72.3%; Score 549; DB 10; Length 774;	
Best Local Similarity		98.1%; Pred. No. 1.7e-148;	
Matches 609; Conservative		0; Mismatches 5; Indels 7; Gaps 5;	
Oy	1	atcttgtagaacaccagtggttgtacctgagaanaaatccccctagatatgatgacaattaac	60
Df	159	ATGTGTAAGCACCCAGTGGGTGTGACCTGGAAGAAATCCCCTGATGATGATGACCTTAAC	218
Oy	61	accatagaattcaaatactctogcctactacaccaagaacatcatgcttccaagaacacccc	120
Df	219	ACCATGGAATTCAAAATTCCTCGCCTACTAGAACACATCATGTGTTCAAGACACCCCC	278
Oy	121	gtctctcttccacaaaagctgctgagaacaagaagtgtgtccagaagggcctgyggaat	180

Db	279	GCTCTTTTCAACCAAGCTGCTGAGAACAAAGAATTTCCTCCAGAGGGGCTGGGAAAT	338
QY	181	tgttcaagcaaatgagatcagatcagagaagtgcattgagccttcgcaaaattccaatcagat	240
Db	339	TTTTTGGCAAAAGAGCAATGAGCAGAGGAGTGTATGGCTTTGGAGAAATTCCTCAATTCAGT	398
QY	241	gagaagggcataaccttggcaagaaaaagctctctcttgaagaagcattctttgagtagtg	300
Db	399	GAGAAAGGCATTAACCTTGCGCAAGAAAAGTCTTCTTGAAAGCATTTCTTTGGAGTAGTG	458
QY	301	gagaaagaaagatctgcagaagcaagcgccttcgcaagctctctgctcagaagtgcataaagaaagt	360
Db	459	GGAAGAGGAAGATTCACAGAGCAAGCGCTGCCAAGGTCTGTGCTCAGGGTCAAAAGAGCTTG	518
QY	361	gaatccaagatctgcagaagccagagagtgatcagtgatgcttcttaagctgaagcagctgc	420
Db	519	GAAATCCAAAGATTCGACACGCCAGCATGTGTCTCAAGTGTCTTTCTTAACCTGAGCAGTGC	578
QY	421	-ttggaagcatgaagctctgtagaccccaagaatcattccattgcgaaccgagtagctgaat	479
Db	579	TTTGGAGCATGAAAGCTGTGGACCCCAAAATCATTTCCATTGGCCAAACCGAGTGTGAAT	638
QY	480	tgattatctctggtccaccacccaagaagcgagccagcagagagcttcgaattccaagaagat	539
Db	639	TGTTTACTCTCTGCG--ACACCAACAAGCCA-CCAGGCGAGGAGGCTTCAATCCAAAGAGAT	695
QY	540	tttttlaactgaaggtctctctcttcacagcttcaa-ggcaacgtgctgtagtctcaagt	598
Db	696	TTTTTAACTGAGGGTCTCTCTCTTCAGTGCCAAAGGCCACGCTGCTGTAG--TTCAATT	753
QY	599	ctaagaagaagatgaagaagaac	619
Db	754	CTAAGCAGATCAAGAAAGAAC	774

FEATURES	SOURCE
LOCUS	B1838169
DEFINITION	B1838169 715 bp mRNA linear EST 04-OCT-2001
ACCESSION	603083677/F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5222875 5',
VERSION	RNA sequence.
KEYWORDS	B1838169
SOURCE	B1838169.1 GI:15949719
ORGANISM	EST.
REFERENCE	human.
AUTHORS	Homo sapiens
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
COMMENT	1 (bases 1 to 715)
	NIH-MGC http://mgs.nci.nih.gov/.
	National Institutes of Health, Mammalian Gene Collection (MGC)
	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.
	Email: cgapds@email.nih.gov
	Tissue Procurement: Life Technologies, Inc.
	cDNA Library Preparation: Life Technologies, Inc.
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNI)
	DNA Sequencing by: Incyte Genomics, Inc.
	Clone distribution: MGC clone distribution information can be
	found through the I.M.A.G.E. Consortium/LMNI at:
	http://image.llnl.gov
	Plate: Llam11560 row: 1 column: 20
	High quality sequence stop: 715.
	Location/Qualifiers
	1..715
	/organism="Homo sapiens"
	/dd_xref="taxon:9606"
	/clone="IMAGE:5222875"
	/clone_id="NIH_MGC_120"
	/lab_host="DH10B"
	/note="Organ: pooled pancreas and spleen; Vector:
	pcmv-sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
	source anonymous pool of spleen and pancreas from 28 yo
	male. Library is oligo-dT primed and directionally cloned

QY	308	aagaatgcgaagacgaccttcgaaggtccttgctcgcaaggtccaaagacgttggaaataacc	367
Db	276	AAGATTGGCAAGACGCGCTCCAAAGGCTTGCTGCATGAAGTCAAAGGACGTGGAAATACC	335
QY	368	aagaattgcacacagcacaagtcggtcccaagtgctctttccaacgttgagacagtcttggagc	427
Db	336	AAGATTGCCACAGCCACGACAGTGTGTCCATGTTCTTCTTAACGTGGAGCAGTGTCTGGAGC	395
QY	428	atgaagctgtgacccccaaagtcatt	453
Db	396	ATGAAGGTGANCATCTGGGATTCCTTT	421

RESULT	8
AA399486	
LOCUS	406 bp mRNA linear EST 16-MAY-1997
DEFINITION	z660c07.t1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:726732
ACCESSION	U000046
	5' / mRNA sequence.

ANIMALISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 406)	
REFERENCE	
AUTHORS	Hillier, L., Allen, M., Bowles, L., Dubouque, T., Geisel, G., Jost, S.,

TITLE
JOURNAL
COMMENT

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel.: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28n13, rev. ET from Amesham.

FEATURES	SOURCE	Location/Qualifiers
1..406		
/organism="Homo sapiens"		
/db_xref="GDB:5923643"		
/db_xref="taxon:9606"		
/clone="IMAGE:726732"		
/clone_1lb="Soares_tests_NHT"		
/sex="male"		
/lab_host="DH10B"		
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5'] TGTTACCAATCTGAAGTGGAGGCGGCCCAATTTTTTTTTTTT 3'). Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."		
127 a	95 c	101 g 82 t 1 others

	Query Match	42.3%	Score 321.4;	DB 9;	Length 406;
	Best Local Similarity	97.7%;	Pred. No. 1.9e-82;		
	Matches 347; Conservative	0;	Mismatches 6;	Indels 2;	Gaps 2;
Oy	1 atgtttagaccagcagggtgtgacctggagaagaacccctcatgatgatgatgaactaac	60			
Db	52 ATGTATACACCACAGCGGTTGCACCTGGAGAAATCCCTCATGATGATGATCACTTAAC	111			
Oy	61 accatagattcaaatcctgcctactacaacagacatcatgtcttcaagagaccct	120			

Db	112	ACCATGAAATTCAAATTCCTCGCTACTACACCGACATCATGCTCTTCACAGACACCCT	171
Qy	121	gcat-tcttcacacaagcgtcgtgagagacaagaatttgcctcagaagggccctgggagaa	179
Db	172	GCTTGCTCTTCACCAAAAGCTGCTGAGACACAAAGATTGTGCCAGAGGGCCATGGGAA	231
Qy	180	ttgttcagacaatgaatcatgagacaagtgctcatggtccttcaga-aattcccaatcca	238
Db	232	TTGTTCAACCAATGATGTCATGACAGAGGTTCATGCGCTTGCAAGAAATTTCCCAATCCA	291
Qy	239	gttgaagaagccataaaccttgcagaagaaagctctcttgcagaaagcattcttggagtag	298
Db	292	GTGGAAGGCCATTAACCTTGCGCAAGAAAGATGTTCTTGGAAGCATCTTGTGGAGTAG	351
Qy	299	ttggaagaagaattctgcagagacagccttgcagaagctctgcgcgaagtgtaagaag	353
Db	352	TGGAGAAGGAAGATTTCGAGAGCAACGCTCGCCAAAGTCTTGCTCAGGGTCAAG	406

RESULT	9
LOCUS	AK016997
DEFINITION	AK016997 1535 bp mRNA linear HTC 19-JAN-2002 Mus musculus adult male testis cDNA, Riken full-length enriched library, clone:933430J01.BC12-11ke apoptosis inhibitor containing protein, full insert sequence.
ACCESSION	AK016997
VERSION	AK016997.1 GI:12856033
KEYWORDS	HTC; CAP trapper.
SOURCE	Mus musculus (strain:C57BL/6J) adult male testis cDNA library clone.11b:RIKEN full-length enriched mouse cDNA library

ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (sites)
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
REFERENCE	2 (sites)

AUTHORS
 Carninci, P., Shibata, K., Hayatsu, N., Suganara, T., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL
 204699374
 MEDLINE
 PUBMED
 11042159
 REFERENCE
 3 (sites)
 AUTHORS
 Shibata, K., Itoh, M., Alzawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujimake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
 TITLE
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 20530913
 MEDLINE
 PUBMED
 11076861
 REFERENCE
 4 (sites)
 AUTHORS
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL
 Nature 409, 685-690 (2001)
 REFERENCE
 5 (bases 1 to 1535)
 AUTHORS
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hitaoka, T., Hori, F.,
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,

SOURCE	Mus musculus (Strain:C57BL/6J) adult male colon cDNA to mRNA, clone.lib.RIKEN full-length enriched mouse cDNA library clone:9030625M01.
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	Carninci,P. and Hayashizaki,Y. 1 (sites)
AUTHORS	Carninci,P. and Hayashizaki,Y. High efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253
JOURNAL MEDLINE	10349636
PUBMED	10349636
REFERENCE	2 (sites)
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374
JOURNAL MEDLINE	11042159
PUBMED	11042159
REFERENCE	3 (sites)
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komono,H., Akiyama,J., Nishi,K., Kitsuana,T., Teshito,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,H., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashwaqi,K., Fujiwako,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kirai,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichipillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913
JOURNAL MEDLINE	11076861
PUBMED	11076861
REFERENCE	4 (sites)
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 665-690 (2001) 5 (bases 1 to 1578)
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL REFERENCE	Nature 409, 665-690 (2001)
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Butt,C., Carandini,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiroka,T., Hoti,F., Hummel,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kuhihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Nunazaki,R., Ono,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Salto,H., Saito,R., Sakai,K., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shitaki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yananaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 220-0045, Japan (E-mail:genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGGAGAAGATCCCAAGACCTCTTTTGTTCCTTTTTCVN 3'], cDNA was prepared by using reverse transcriptase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.2. Second strand cDNA was prepared with the primer

FEATURES	source
gagaggt of sequence [5', GAGAGGACATATCTCCAGTTAATTAATTAAATCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.	
Location/Qualifiers	1..1578
/organism="Mus musculus"	
/strain="C57BL/6J"	
/db_xref="MGD:MGI:1894816"	
/db_xref="taxon:10090"	
/clone="9030625M01"	
/sex="male"	
/tissue="colon"	
/clone_id="RIKEN full-length enriched mouse cDNA library"	
/dev_stage="adult"	216..1202
/note="evidence:NAS hypothetical protein putative"	
/codon_start=1	
/protein_id="BAB31290.1"	
/db_xref="GI:12858358"	
/translation="MCSYVDLEDIDPLEDDPNSTIEFKILAFARHHVKNPAPVS PKLSRPSLSOKALGTSYDPSMNOVSIPCGSPSEKNIISLGKSSMRLLFPAEKE EGLPSPKREIRAGCPQGPFPYEROSVGHNDHNPMSLSVEDRLESYDVKATIANR VAELVSWPPDPYIHSQGSGLKEREVEILYFREGGCDSKNKGEODIISKIVELK FSGQGLGREIKKDKALMSFQDGLSYSTFTIDPLRDVDTRESEVKKRGFMAALA IDALAKRLAIDNHNMRLMGLFGYKLYREYFSPVQONGWEKILGISHEVD"	
BASE COUNT	427 a 393 c 420 g 338 t
ORIGIN	
Query Match	38.8%; Score 294.6; DB 11; Length 1578;
Best Local Similarity	69.6%; Pred. No. 1.9e-74;
Matches	482; Conservative 0; Mismatches 184; Indels 27; Gaps 5;
QY	1 atgtgtagcaccagttgggtgtgtgacctgtgaagaatcccccagatgatgtgacctaac 60
DB	216 ATGTGACACACACAGTGTTATGACTCGGAAGACATCCCTGAGAGATGATCTCAAC 275
QY	61 accatagattcaaatctctgcgcctacacaccgaactatgttcttaagagccccc 120
DB	276 AGCATGAGTTCAAAATCTTGCGCTTCTTACGCCGACACACCATGTTTCAAGAACACCCG 335
QY	121 gctctcttcaccaagaagctgtgagaacaaagtgtgtcccaagaggccttgaggat 180
DB	336 GCTGCTTCTTGCCCCAACACTCTCCGAACMAGAGTCTGTCCAGAAAGCCCTGGGACT 395
QY	181 tgttcagaatagtcatgtgacagagtgatcatgacctgtgcagaattcccaatcagt 240
DB	396 TGGT---CAACTGATCTCTGACACAGATATCATGTGCTTGACAGAGTTTCCCTCCAGC 452
QY	241 gagaagagccataaaccttggccaagaanaagctcttcttgaaagactctttggaatagt 300
DB	453 GAAAGAACATTCAGCTTGCGGCAAGAAAGATCTTGTGGAACACTCTTCAAGGTGGCC 512
QY	301 gaga---aggagaatttcgacagagcagcgcctgcgaagtcctctgcaggttcaagaagc 357
DB	513 GAGAAAGAGAAAGGCTGCGCAAGCTCCCAAGAGAGATTCGAGCTCAGGGTCTCTAGGC 572
QY	358 ttggaataccaagaattcg-----cacagccagcagtggttcagagtgtcttct 405
DB	573 CCGTTCGCCGTTAGACGGGACAGTGGCTTCCACACACACAGCACTGGCCACAGGCTCTGAGC 632
QY	406 aacgtgagcagtgctgttgagcatgaagctgtgtgacccccaagtatcttcattcgcaac 465
DB	633 AGTGTGGAGCAGCGGCTTGAGAGTGAATGTGGATTGCCAAAGTGGCTTGATTGCCAAC 692
QY	466 cgaagtgcgtgaatgtttatctcctgagccacacacacaaagcgagcccaaggcaggaagcttc 525
DB	693 AGAGTGGCTGAATTTGTTTACTCTCTGGCCACCCACCGAGATGTTCATCCACAGCAGGGAAGA 752


```

Db 370 GCTGCTTCTCCGCAAGCTCTCCAGAACAGAGAGTCTGCCAGAAAGCCCTGGGACT 429
Oy 181 tcttcagaataatgaatcagaagagtgatcattgcttcgcaagaattcccaatcagt 240
Db 430 TGGT---CACTGATTTCTCGGACACAGATATCTGCTTGCAGAGGTCTCCCTCCAGC 486
Oy 241 gagaagccataaacttgcgaagaaagtctctctggaagacatctcttgagagtagtg 300
Db 487 GAAAGAACATCAGCTTGGGCAAGAGAGTCTTCTTGAGAACACTCTTCAGGCTGCC 546
Oy 301 gaga---aggaagatcgcagagcagcctgcgaaggtctctgctcagagtgcaagagc 357
Db 547 GAGAGGAGGAGAGGAGGCTGCCAGCTCCCAAGAGAGATCGAGCTCAGGGTCTCAGGGC 606
Oy 358 ttggaataccaaagtctg-----caagccagagagtggtccaaagtctcttct 405
Db 607 CCGTTCCTCCGCTGAGACCGGAGAGTGGCTTCCACACACGACGCTGGCCAGGTCTTGAGC 666
Oy 406 aacgtgagcagtgcttggaagcatgaagctgtgagcccaagtcattccattgccaac 465
Db 667 AGTGTGGAGCAGCCGCTGAGAGTGAAGTGTGGATTCCAAAGTGGCTTGATTGCCAAC 726
Oy 466 cgaatagctgaatgttattctctgcccacacacacagcagcagcagagagagcttc 525
Db 727 AGAGTGGCTGAATGTCTTACTCTGCGCACACAGATGTCTACACAGCCAGGAGAGA 786
Oy 526 aagtcacaaagatttctgaatcagaggtctctctccctcagctcagagcagagctc 585
Db 787 AGCAGAGCTCAAGAGAGGCTCTCGGAGATCTGTACTTCAAGTTGAGAGAC---CTT 841
Oy 586 gtagcttcaagttctaaagaatgaagaagaacataactgaacaaatgttgaagctg 645
Db 842 GTGACTCTTAAG---AATTAAGATGAGAGAACCAATATAGCAAGATTGTTGAGACTG 897
Oy 646 ctgaatatctcagagatcagtgatggaagaag 678
Db 898 CTGAATTTCTCGGGGATCAGTTGGGAAGAGAG 930

```

```

RESULT 13
BI463753 776 bp mRNA linear EST 21-AUG-2001
LOCUS 603203445F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5269331 5',
DEFINITION mRNA sequence.
ACCESSION BI463753
VERSION BI463753.1 GI:15254409
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
1 (bases 1 to 776)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaapb-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L16M1679 row: 1 column: 12
High quality sequence stop: 685.
Location/Qualifiers
1. 776
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5269331"

```

FEATURES

source

```

/clone.lib="NIH_MGC_97"
/lab_host="DH10B"
/node="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTVN-3',
size-selected for average insert size 2.2 kb and
normalized to 10^5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 243 a 180 c 191 g 162 t
ORIGIN

```

```

Query Match 36.4%; Score 276.2; DB 10; Length 776;
Best Local Similarity 91.2%; Pred. No. 3.3e-69;
Matches 340; Conservative 0; Mismatches 23; Indels 10; Gaps 4;

```

```

Oy 1 atgtgtgagccagtggtgtgtgacctggaagaatccccatagatgatgaactaac 60
Db 400 ATGTGTGAGCACCAAGTGGGTGTGACCTGGAGAAATCCCTAGATGATGATCACTAAAC 459
Oy 61 accatgaattcaaatccctcgactactacacagacatcatgtcttcaagagcaccct 120
Db 460 ACCATGAATTCATAAATCCTCGCTTACTACACGACATCATGTCTTCAAGAGCACCCCT 519
Oy 121 gctctcttcacaaagctgtctgagaacaaagattgtctccagaaggcctgaggaaat 180
Db 520 GCTCTCTTCTCACCAAAAGCTGTGAGAAAGAGATTGTGCCAGAGGGGCTGGGGAAAT 579
Oy 181 tcttcagcaatgaatcagatgagagagtgatc-ggctctgcagaaatccccatcag 239
Db 580 TCTTTCAGCAATATGATGATGAGAGAGGTGTCAATGGGCTTGCAGAAATTCCTCAATCCAG 639
Oy 240 tgaaga-ggcccataaaccttgcgaaga---agctctcttgcgaagcattcttggag 295
Db 640 TGAAGACGGCCATTAACCTTGGCAACGAAACAGAGTCTTCTTGGAAAGCCATTTCTTGG 699
Oy 296 tagt-----ggagaaggaagatcgcagagcagcgtgcgaaggtctctgcagagta 350
Db 700 CAGTACTTGGAGAACCGAATTCGAGAGACGCGCTGCCAAAGCTCTGTGTCAGAGGTCA 759
Oy 351 aaggaagcttgga 363
Db 760 AAGGACCGTGGGA 772

```

```

RESULT 14
BI453888 920 bp mRNA linear EST 21-AUG-2001
LOCUS 603174765F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5253808 5',
DEFINITION mRNa sequence.
ACCESSION BI453888
VERSION BI453888.1 GI:15244544
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 920)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaapb-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2002, 05:05:29 ; Search time 302.61 Seconds
(without alignments)
4306.328 Million cell updates/sec

Title: US-09-771-961-3
Perfect score: 759
Sequence: 1 atgtgtagaccagttggtg.....gtttgctgctgatttga 759

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759	100.0	759	22	AA013236
2	759	100.0	2132	22	AA013237
3	757.4	99.8	954	22	AAH22583
4	678	89.3	984	22	AA013235
5	676.4	89.1	1179	22	AAH22582
6	485.6	64.0	1917	23	AA050453
7	429	56.5	632	22	AA01348
8	388	51.1	388	22	ABA47100
9	388	51.1	388	22	ABA64982

10	388	51.1	388	22	ABA32090	Probe #10556 for g
11	388	51.1	388	22	AAK13408	Human brain expres
12	388	51.1	388	22	AAK39145	Human bone marrow
13	388	51.1	388	22	AAI19954	Probe #9887 for ge
14	388	51.1	388	22	AAI45152	Probe #13838 used
15	388	51.1	388	22	AAI05666	Probe #5657 used t
16	388	51.1	388	22	ABA52358	Human foetal liver
17	317	41.8	466	22	ABA22153	Probe #619 for gen
18	317	41.8	466	22	AAK00629	Human brain expres
19	317	41.8	466	22	AAK26079	Human bone marrow
20	317	41.8	466	22	AAI10708	Probe #641 for gen
21	317	41.8	466	22	AAI31966	Probe #652 used to
22	317	41.8	466	22	AAI00638	Probe #629 used to
23	274.4	36.2	987	22	AAK22670	Mouse Bcl-G polype
24	83	10.9	8922	22	AAH87069	Human Bcl-G polype
25	39.8	5.2	43	22	AAH22593	Human Bcl-G polype
26	39.8	5.2	43	22	AAH22594	Human Bcl-G polype
27	37	4.9	7055	20	AAH22594	Bcl-6s mutagenic p
28	37	4.9	7783	20	AAH26302	Sequence of phage
29	36.2	4.8	954	22	AAH22583	Human Bcl-X-like p
30	36.2	4.8	954	22	AAH22583	Human Bcl-X-like p
31	36.2	4.8	984	22	AAH22583	Human Bcl-X-like p
32	36.2	4.8	1179	22	AAH22582	Human Bcl-X-like p
33	36.2	4.8	2132	22	AAH22582	Human Bcl-X-like p
34	35.8	4.7	10377	24	ABU34220	Human Bcl-X-like p
35	35.2	4.6	4019	21	AAZ43878	Human immune syste
36	34.6	4.5	6971	20	AAZ6304	C. vicina arylphor
37	34.2	4.5	600	20	AAV86185	Sequence of phage
38	34.2	4.5	1001	21	AAH51276	EST clone J635. H
39	34	4.5	917	22	AAI94362	Human GSHS related
40	34	4.5	32195	22	AAI94362	Human neuroblastom
41	34	4.5	32219	22	AAI94362	Human neuroblastom
42	33.8	4.5	3280	18	AAH87981	Human cardiovascular
43	33.8	4.5	2755	18	AAH87981	Human cardiovascular
44	33.8	4.5	3225	22	AAH87981	PCAO-1 modular ve
45	33.8	4.5	3488	18	AAH87982	PCAO-1 modular ve

ALIGNMENTS

RESULT 1	
AA013236	standard; cDNA; 759 BP.
ID	AA013236
XX	AA013236
AC	AA013236
XX	23-OCT-2001 (first entry)
XX	Human BCL-X-like protein encoding cDNA #2.
XX	Human; BCL-X-like protein; therapy; physiological disorder; ss.
OS	Homo sapiens.
XX	
XX	Key
XX	Location/Qualifiers
XX	1..759
XX	/*tag= a
XX	/product= "Human BCL-X-like protein #2"
XX	
XX	WO200157213-A2.
XX	09-AUG-2001.
XX	
XX	31-JAN-2001; 2001WO-US03446.
XX	
XX	04-FEB-2000; 2000US-0180412.
XX	
XX	(LEXI-) LEXICON GENETICS INC.
XX	
XX	Donoho G, Hilbun E, Turner CA, Friedrich G, Abulin A, Zambrowicz B,
XX	Sands AT;
XX	

DR WPI; 2001-488882/53.
DR P-PSDB; AAE07041.

DR P-PSDB; AAEO7041.

PT Novel isolated human BCL-X-like polynucleotide, useful in therapeutic, diagnostic and pharmacogenic applications -

PS Disclosure; Page 31; 33pp; English.

The present sequence is a cDNA encoding human BCL-X-like protein. The BCL-X-like polynucleotides are useful in the therapeutic, diagnostic and pharmacocentric applications. They are useful for screening drugs effective in the treatment of symptomatic or phenotypic manifestations perturbing the normal function of protein in the body and also for treating physiological disorders and diseases. The BCL-X-like polynucleotides are useful in conjunction with polymerase chain reaction to screen libraries, isolate clones, to prepare cloning and sequencing templates and as hybridisation probes for assessing gene expression patterns.

SQ Sequence 759 BP; 216 A; 182 C; 185 G; 176 T; 0 other;

Only Match	100.0%	Score 759;	DB 22	length 759;
Best Local Similarity	100.0%	Pred. No. 2.7e-228;		
Matches 759; Conservative	0;	Mismatches	0;	Gaps 0

Qy	1	atggtgacacacagtggtgtgtgcacgtcggaagaatccccataatgatgatgacccaac	60
Db	1	atggtgtgcacccacgtgggtgtgtacccctggagaatccccctatgatatgtgtgccaaac	60
Qy	61	accatagaattccaataatcctcgtcctactaaccagaacatcatgtcttcaagagaccctc	120
Db	61	accatagaattccaataatcctcgtcctactaaccagaacatcatgtcttcaagagaccctc	120
Qy	121	gtctctctccacccaagcgtcgtgagaacaagaattgtcccaagaggggccggggaat	180
Db	121	gtctctctccacccaagcgtcgtgagaacaagaattgtcccaagaggggccggggaat	180
Qy	181	tgttcagcaaaatgattcatctatgacagagtgatcatgtgccttcgcagaatccccatccagt	240
Db	181	tgttcagcaaaatgattcatctatgacagagtgatcatgtgccttcgcagaatccccatccagt	240
Qy	241	gagaaagcgccataacctgtgcagaagaaaagtctcttcttggaagacattctttgagtcagt	300
Db	241	gagaaagcgccataacctgtgcagaagaaaagtctcttcttggaagacattctttgagtcagt	300
Qy	301	gagaagagaagaattcgcagagcagcgcctgcacaaggtctctctcagtggtccaagaagcgttg	360
Db	301	gagaagagaagaattcgcagagcagcgcctgcacaaggtctctctcagtggtccaagaagcgttg	360
Qy	361	gaataccaagaattcgcacagccagtcagtgctcagtgctcttcttaacgttgagacagtcg	420
Db	361	gaataccaagaattcgcacagccagtcagtgctcagtgctcttcttaacgttgagacagtcg	420
Qy	421	tttgagacatgaacgtgtgagaccccaagatatttccatcttcacgcacagatgattgaaat	480
Db	421	tttgagacatgaacgtgtgagaccccaagatatttccatcttcacgcacagatgattgaaat	480
Qy	481	gttatctctgcgcacccacacacaagcgaccacagcgagaggttcaagtccaagaagatt	540
Db	481	gttatctctgcgcacccacacacaagcgaccacagcgagaggttcaagtccaagaagatt	540
Qy	541	tttgtaactgagggtctctctcttcacgtctccaagcgcaagtgctgtgattgaatttct	600
Db	541	tttgtaactgagggtctctctcttcacgtctccaagcgcaagtgctgtgattgaatttct	600
Qy	601	aagaagaatgaagaagaacaataactgcgcaaaattgttagctgtcgaataatctcaaga	660
Db	601	aagaagaatgaagaagaacaataactgacccaataattgttagctgtcgaataatctcaaga	660
Qy	661	gatacagtttgaaagaagaagacatgcctcatatcccatctccctgtgtgacacacagcatc	720
Db	661	gatacagtttgaaagaagaagacatgcctcatatcccatctccctgtgtgacacacagcatc	720

```
QY      721 caaggtttccacagaatggttgatggcctgcatttga 755  
       |||||  
Db      721 caagggtttccacagaatggttgatggcctgcatttga 755
```

RESULT 2

ID AAD13237 standard; cDNA; 2132 BP

AC AAD13237;

DT 23-OCT-2001 (first entry)

Human BCL-X-like protein encoding cDNA #3.

KW Human; BCL-X-like protein; therapy; physiological disorder; ss.

OS Homo sapiens

PN W0200157213-A2

PD 09-AUG-2001

31-JAN-2001; 2001WO-US03446.

PR 04-FEB-2000; 2000US-0180412.

PA
XX (LEXI-) LEXICON GENETICS INC.

PI Donoho G, Hilbun E, Turner CA, Friedrich G, Abuln A, Zambrowicz B,
PI Gande AM:

DR WPI; 2001-488882/53

PT Novel isolated human BCL-X-like polynucleotide, useful in therapeutic
PT diagnostic and pharmacogenic applications -

PS Disclosure; Page 32-33; 33pp; English.

CC The present sequence is a cDNA encoding human BCL-X-like protein.
CC
CC The BCL-X-like polynucleotides are useful in therapeutic, diagnostic
CC and pharmacogenic applications. They are useful for screening drugs
CC effective in the treatment of symptomatic or phenotypic manifestations
CC perturbing the normal function of protein in the body and also for
CC treating physiological disorders and diseases. The BCL-X-like
CC polynucleotides are useful in conjunction with polymerase chain
CC reaction to screen libraries, isolate clones, to prepare cloning
CC and sequencing templates and as hybridisation probes for assessing
CC gene expression patterns.

SQ Sequence 2132 BP; 615 A; 477 C; 486 G; 549 T; 5 other,

Query Match	100.0%	Score 759	DB 22	Length 2132
Best Local Similarity	100.0%	Pred. NO	4.6e-228	
Matches 759	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Oy	1	atgctgtagacacagttggtgtgtgtgaccccttggaagaatcccccatagtatgtatgacccaac	60
Db	192	atgtgttagaccacagctgggtgtgtgaccccttggaagaatccccctagtatgtatgacccaac	251
Oy	61	accatagaaattccaatccctcgtgctactacaaccgaacatcatgtcttcaagaagacccc	120
Db	252	accatagaaattccaatccctcgtgctactacaaccgaacatcatgtcttcaagaagacccc	311
Oy	121	gctcctctctccacaagaactgtctgagaaacaagaattgttcccaagagggccttgggaat	180
Db	312	gctcctctctccacaagaactgtctgagaaacaagaattgttcccaagagggccttgggaat	371
Oy	181	tgttcagaagaattggtcatctgtaacagaagttctcatagccttggcagaattcccaatccagt	240
Db	372	tgttcagaagaattggtcatctgtaacagaagttctcatagccttggcagaattcccaatccagt	431

QY 241 gagaagccataaaccttggcaagaagaagtctcttggaaagcattcttggatgag 300
DB 432 gagaagccataaaccttggcaagaagaagtctcttggaaagcattcttggatgag 491
QY 301 gagaaggaagaatctgcagagacagcctgcgcgaagtctcttgcaggttcaaaaggagcttg 360
DB 492 gagaaggaagaatctgcagagacagcctgcgcgaagtctcttgcaggttcaaaaggagcttg 551
QY 361 gaatacaagaatctgcagagacagcctgcgcgaagtctcttgcaggttcaaaaggagcttg 420
DB 552 gaatacaagaatctgcagagacagcctgcgcgaagtctcttgcaggttcaaaaggagcttg 611
QY 421 ttggagcatgaagctgtgagaccccaaaagcttccatctgcgaacccgaagtagctgaatt 480
DB 612 ttggagcatgaagctgtgagaccccaaaagcttccatctgcgaacccgaagtagctgaatt 671
QY 481 gttattctctgcgcacacacacacagcagccagcagagagcttcaagtccaaagagatt 540
DB 672 gttattctctgcgcacacacacacagcagccagcagagagcttcaagtccaaagagatt 731
QY 541 ttgttaactgaggtctctctctccacagctcccaagcagctgctgtaagcttcaagttct 600
DB 732 ttgttaactgaggtctctctctccacagctcccaagcagctgctgtaagcttcaagttct 791
QY 601 aagaagaatgaagaagaacaataactagccaagaattgttgaagctgtgtaataatcaga 660
DB 792 aagaagaatgaagaagaacaataactagccaagaattgttgaagctgtgtaataatcaga 851
QY 661 gatcaattgtgaaagaagaagacactgtccttccatcccttgcctgtgttgaaaccagcattc 720
DB 852 gatcaattgtgaaagaagaagacactgtccttccatcccttgcctgtgttgaaaccagcattc 911
QY 721 caagggtttccacagagatgttgatgagcctgcatcttga 759
DB 912 caagggtttccacagagatgttgatgagcctgcatcttga 950

RESULT 3
AAH22583
ID AAH22583 standard; cDNA; 954 BP.
AC AAH22583;
XX
DT 07-SEP-2001 (first entry)
XX
DE Human Bcl-GS polypeptide encoding cDNA.
XX
KW Bcl-G; cancer; cancer therapy; oncogene; apoptosis; Bcl-Gs; cytosolic;
XX antiapoptotic; chromosome 12p12.3; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 196..954
FT /tag= a
FT /product= "Bcl-GS"
XX
XX MO200144282-A2.
XX
XX PD 21-JUN-2001.
XX
XX PF 13-DEC-2000; 2000MO-US33793.
XX
XX PR 14-DEC-1999; 99US-0461641.
XX
XX PA (BURN-) BURNHAM INST.
XX
XX PI Reed JC, Godzik A;
XX
XX DR WPI: 2001-398125/42.
XX P-PSDB: AAB85167.
XX

PT Novel polynucleotide encoding Bcl-G polypeptide, useful for modulating
PT apoptosis, and for diagnosing and treating cancer
XX
XX
PS Claim 3; Fig 3; 117pp; English.
XX
CC The invention relates to Bcl-G polypeptides and nucleic acids encoding
CC them. The Bcl-G polypeptides can be expressed by standard recombinant
CC methodology. Bcl-G oligonucleotides (or its anti-sense strand) and Bcl-G
CC specific antibodies are useful for diagnosing cancer, monitoring cancer
CC therapy or assessing prognosis of patients with cancer. The Bcl-G
CC polypeptides are useful for modulating the activity of an oncogenic
CC polypeptide. They are useful for identifying modulators, for modulating
CC a level of apoptosis mediated by the Bcl-G polypeptide. A therapeutic
CC composition comprising the Bcl-G polypeptide, polynucleotide or antibody
CC is useful for treating a pathology characterized by abnormal cell
CC proliferation especially cancer. The present sequence represents a cDNA
CC encoding a Bcl-GS polypeptide.
XX
SQ Sequence 954 BP; 267 A; 237 C; 240 G; 210 T; 0 other;

Query Match 99.8%; Score 757.4; DB 22; Length 954;
Best Local Similarity 99.9%; Pred. No. 9,8e-228;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgttaacacacagtggtgtgacactggaagaatccacctagatgtagtacctaac 60
DB 196 atgttaacacacagtggtgtgacactggaagaatccacctagatgtagtacctaac 255
QY 61 accatagaattcaaaaatccctgcctactacacacagatcatgtcttcaagaaccct 120
DB 256 accatagaattcaaaaatccctgcctactacacacagatcatgtcttcaagaaccct 315
QY 121 gctcctctcccaagaagtgtgtgaaagaagaatttgcctccagaaggccttgggaat 180
DB 316 gctcctctcccaagaagtgtgtgaaagaagaatttgcctccagaaggccttgggaat 375
QY 181 ttttaagcaaatgaatcattgacagagaggtgtcattgagccttgcagaattcccaatccagt 240
DB 376 ttttaagcaaatgaatcattgacagagagaggtgtcattgagccttgcagaattcccaatccagt 435
QY 241 gagaagccataaaccttggcaagaagaagtcttcttggaaagcattcttggatgag 300
DB 436 gagaagccataaaccttggcaagaagaagtcttcttggaaagcattcttggatgag 495
QY 301 gagaaggaagaatctgcagagacagcctgcgcgaagtcttgcaggttcaaaaggagcttg 360
DB 496 gagaaggaagaatctgcagagacagcctgcgcgaagtcttgcaggttcaaaaggagcttg 555
QY 361 gaatacaagaatctgcagagacagcctgcgcgaagtcttgcaggttcaaaaggagcttg 420
DB 556 gaatacaagaatctgcagagacagcctgcgcgaagtcttgcaggttcaaaaggagcttg 615
QY 421 ttggagcatgaagctgtgagaccccaaaagcttccatctgcgaacccgaagtagctgaatt 480
DB 616 ttggagcatgaagctgtgagaccccaaaagcttccatctgcgaacccgaagtagctgaatt 675
QY 481 gttattctctgcgcacacacacacagcagccagcagagagcttcaagtccaaagagatt 540
DB 676 gttattctctgcgcacacacacacagcagccagcagagagcttcaagtccaaagagatt 735
QY 541 ttgttaactgaggtctctctctccacagctcccaagcagctgctgtagcttcaagttct 600
DB 736 ttgttaactgaggtctctctctccacagctcccaagcagctgctgtagcttcaagttct 795
QY 601 aagaagaatgaagaagaacaataactagccaagaattgttgaagctgtgtaataatcaga 660
DB 796 aagaagaatgaagaagaacaataactagccaagaattgttgaagctgtgtaataatcaga 855
QY 661 gatcaattgtgaaagaagaagacactgtccttccatcccttgcctgtgttgaaaccagcattc 720
DB 856 gatcaattgtgaaagaagaagacactgtccttccatcccttgcctgtgttgaaaccagcattc 915

QY 721 cagggtttccacagagatgtttgatgagcctgacttga 759
DB 916 cagggtttccacagagatgtttgatgagcctgacttga 954

RESULT 4
ID AAD13235 standard; cDNA; 984 BP.
AC AAD13235;
DT 23-OCT-2001 (first entry)
XX Human BCL-X-like protein encoding cDNA #1.
DE Human; BCL-X-like protein; therapy: physiological disorder; ss.
KW Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 1..984
F1 /product= "Human BCL-X-like protein #1"

XX MO200157213-A2.
XX 09-AUG-2001.
XX 31-JAN-2001; 2001WO-US03446.
XX 04-FEB-2000; 2000US-0180412.
XX (LEXI-) LEXICON GENETICS INC.
PI Donoho G, Hilbun E, Turner CA, Friedrich G, Abuin A, Zambrowicz B;
PI Sands AT;
XX WPI: 2001-488882/53.
XX P-PSDB; AAE07040.
XX Novel isolated human BCL-X-like polynucleotide, useful in therapeutic,
XX diagnostic and pharmacogenic applications -
XX
XX Claim 2a; Page 30; 33pp; English.
XX
XX The present sequence is a cDNA encoding human BCL-X-like protein.
XX The BCL-X-like polynucleotides are useful in therapeutic, diagnostic
XX and pharmacogenic applications. They are useful for screening drugs
XX effective in the treatment of symptomatic or phenotypic manifestations
XX pertaining to the normal function of protein in the body and also for
XX treating physiological disorders and diseases. The BCL-X-like
XX polynucleotides are useful in conjunction with polymerase chain
XX reaction to screen libraries, isolate clones, to prepare cloning
XX and sequencing templates and as hybridisation probes for assessing
XX gene expression patterns.
XX
XX Sequence 984 BP; 289 A; 230 C; 250 G; 215 T; 0 other;

Query Match 89.3%; Score 678; DB 22; Length 984;
Best Local Similarity 100.0%; Pred. No. 9,4e-203;
Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgttagcaccagtggtgtgacctggaagaatccccctagatgatgacttaac 60
DB 1 atgtgttagcaccagtggtgtgacctggaagaatccccctagatgatgacttaac 60
QY 61 accatagaattcaaatctcgcctactacacacagacatcatgtcttaagaagaccct 120
DB 61 accatagaattcaaatctcgcctactacacacagacatcatgtcttaagaagaccct 120
QY 121 gctctcttccacaaagctgctgagacagaagttgtccagagggcctgggaat 180
DB 121 gctctcttccacaaagctgctgagacagaagttgtccagagggcctgggaat 180

DB 121 gctctcttccacaaagctgctgagacagaagttgtccagagggcctgggaat 180
QY 181 tgtcagaacaatgagtcacatgagacagagggtgtcatagccttgacagaattcccaatcagt 240
DB 181 tgtcagaacaatgagtcacatgagacagagggtgtcatagccttgacagaattcccaatcagt 240
QY 241 gagaagccataaacccttgacagaagaagaagctctcttgaaagcattcttgagtagtg 300
DB 241 gagaagccataaacccttgacagaagaagaagctctcttgaaagcattcttgagtagtg 300
QY 301 gagaaggaagaatcgcacagacgccttcgaaggtcctcctcaggttcaaaagacgttg 360
DB 301 gagaaggaagaatcgcacagacgccttcgaaggtcctcctcaggttcaaaagacgttg 360
QY 361 gaataccaagaatcgcacagacgccttcgaaggtcctcctcaggttcaaaagacgttg 420
DB 361 gaataccaagaatcgcacagacgccttcgaaggtcctcctcaggttcaaaagacgttg 420
QY 421 ttggaagcatgaagctgtgagaccacaaagtcatcttcattgccaaccgagttagctgaatt 480
DB 421 ttggaagcatgaagctgtgagaccacaaagtcatcttcattgccaaccgagttagctgaatt 480
QY 481 gttattccttgccacccacaaagcgaccccgagggaggttcaagtcacaaagagatt 540
DB 481 gttattccttgccacccacaaagcgaccccgagggaggttcaagtcacaaagagatt 540
QY 541 ttgttaactgaggggtctctccttcacagtcacaaagcgacgttgacttaagttct 600
DB 541 ttgttaactgaggggtctctccttcacagtcacaaagcgacgttgacttaagttct 600
QY 601 aagaagaatgagaagaacaataactagccaataattgttgaagctgtgtaataattcaaga 660
DB 601 aagaagaatgagaagaacaataactagccaataattgttgaagctgtgtaataattcaaga 660
QY 661 gatcagttggaagaagaag 678
DB 661 gatcagttggaagaagaag 678

RESULT 5
ID AAH22582 standard; cDNA; 1179 BP.
AC AAH22582;
DT 07-SEP-2001 (first entry)
XX Human Bcl-G1 polypeptide encoding cDNA.
DE Human Bcl-G1 polypeptide encoding cDNA.
KW Bcl-G; cancer; cancer therapy; oncogene; apoptosis; Bcl-G1; cytostatic;
KW antiapoptotic; chromosome 12p12.3; human; ss.
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 196..1179.
FT /*tag= a
F1 /product= "Bcl-G1"

XX MO200144282-A2.
XX 21-JUN-2001.
XX
XX 13-DEC-2000; 2000WO-US33793.
XX 14-DEC-1999; 99US-0461641.
XX (BURN-) BURNHAM INST.
XX Reed JC, Godzik A;
XX
XX WPI: 2001-398125/42.
XX P-PSDB; AAB85166.

Oy	60	cacacataaatccaatccctcgcccaccacacagacat--catgtcttcaagaagcacc	117
Db	156	caccatagaattccaatatccctgcctactaccacagacatcatgtcttccaagaagcacc	215
Oy	118	cctgctct-ctctcacccaagaagtgtcttgaaacaagaatttgtlccccagaagggtctggg	176
Db	216	cctgcttcgtcttctccaccaagaagtgtgtgaaacaagaatttgtlccccagaagggtccatggg	275
Oy	177	gaattgttcagaataatgatgtcatgtgacagaagtggtcatgtgaccttgaagaattcccatc	236
Db	276	gaattgttcaagaataatgatgtcatgtgacagaagtggtcatgtgaccttgaagaattcccatc	335
Oy	237	cagttgagaagcgcatanaaaccttggccaagaanaaaagtctctcttgnaaaagcatcttggagt	296
Db	336	cagttgagaagcgcatanaaaccttggccaagaanaaaagtctctcttgnaaaagcatcttggagt	395
Oy	297	agtgtgagaagaagatctgcagagacgacgcttcccaagtgctctgtctcaggtctaaagagac	356
Db	396	agtgtgagaagaagatctgcagagacgacgcttcccaagtgctctgtctcaggtctaaagagac	455
Oy	357	gtctgtaataccaagaattgcacagcgacagtggtctcaggtgtcttctctaagcttgagca	416
Db	456	gtctgtaataccaagaattgcacagcgacagtggtctcaggtgtcttctctaagcttgagca	515
Oy	417	gtgtcttgagacatgaa-----	432
Db	516	gtgtcttgagacatgaaagtgctgtacaaaagccagaagtgcccttagagatttgacacacc	575
Oy	433	-----	432
Db	576	aacctctgtgccaaaaagcccaactcactactgtctcctaagttccctgtgaattactcatagt	635
Oy	433	-----gctgtggaccaccaagatcatlttccatttgcaccacgagtagc	473
Db	636	tagtcttcaaatgttiactatgtctgttggaacccaagaatcatlttccatttgcaccacgagtagc	695
Oy	474	tgaatttgtttattctctgtgcccacacccaagaaggaacccaagggcaggagcttcaaagttcaa	533
Db	696	tgaatttgtttattctctgtgcccacacccaagaaggaacccaagggcaggagcttcaaagttcaa	755
Oy	534	agaagatttttgaactgaggtgtctctccttcacagctccaaggccaagtgctgtgagcttc	593
Db	756	agaagatttttgaactgaggtgtgtctctccttcacagctccaaggccaagtgctgtgagcttc	815
Oy	594	aagttcttaagaagaatgatgaagaagaacaataatactag--ccaatattgttgagctgtcgaat	652
Db	816	aagttcttaagaagaatgatgaagaagaacaataatactagccaataattgttgagctgtcgaat	875
Oy	653	attcaggagatcatgttggaagaagaag	678
Db	876	attcaggagatcatgttggaagaagaag	901
RESULT 7			
ID	AAL01348	AAL01348 standard; cDNA: 632 BP.	
XX	AAL01348:		
AC			
XX			
DT	21-NOV-2001	(First entry)	
XX			
DE	Human reproductive system related antigen cDNA SPQ ID NO: 1349.		
XX			
KM	Human; reproductive system related antigen; reproductive system disorder;		
XX	Cancer; gene therapy; ss.		
OS	Homo sapiens.		
XX			
PN	WO200155320-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-USO1339.		

PR	31-JAN-2000;	2000US-0179065;
PR	04-FEB-2000;	2000US-0180628;
PR	24-FEB-2000;	2000US-0186664;
PR	02-MAR-2000;	2000US-0186350;
PR	16-MAR-2000;	2000US-0189874;
PR	17-MAR-2000;	2000US-0190976;
PR	18-APR-2000;	2000US-0196123;
PR	19-MAY-2000;	2000US-0205151;
PR	07-JUN-2000;	2000US-0209467;
PR	28-JUN-2000;	2000US-0214866;
PR	30-JUN-2000;	2000US-0213135;
PR	07-JUL-2000;	2000US-0216647;
PR	07-JUL-2000;	2000US-0216880;
PR	11-JUL-2000;	2000US-0217487;
PR	11-JUL-2000;	2000US-0217486;
PR	14-JUL-2000;	2000US-0218290;
PR	26-JUL-2000;	2000US-0220963;
PR	26-JUL-2000;	2000US-0220964;
PR	14-AUG-2000;	2000US-0224518;
PR	14-AUG-2000;	2000US-0224519;
PR	14-AUG-2000;	2000US-0225213;
PR	14-AUG-2000;	2000US-0225214;
PR	14-AUG-2000;	2000US-0225266;
PR	14-AUG-2000;	2000US-0225267;
PR	14-AUG-2000;	2000US-0225268;
PR	14-AUG-2000;	2000US-0225270;
PR	14-AUG-2000;	2000US-0225447;
PR	14-AUG-2000;	2000US-0225757;
PR	14-AUG-2000;	2000US-0225758;
PR	14-AUG-2000;	2000US-0225759;
PR	14-AUG-2000;	2000US-0226275;
PR	14-AUG-2000;	2000US-0226681;
PR	22-AUG-2000;	2000US-0226182;
PR	22-AUG-2000;	2000US-0226182;
PR	22-AUG-2000;	2000US-0227009;
PR	23-AUG-2000;	2000US-0228924;
PR	30-AUG-2000;	2000US-0229287;
PR	01-SEP-2000;	2000US-0229287;
PR	01-SEP-2000;	2000US-0229344;
PR	01-SEP-2000;	2000US-0229344;
PR	05-SEP-2000;	2000US-0229309;
PR	05-SEP-2000;	2000US-0229513;
PR	05-SEP-2000;	2000US-0230437;
PR	06-SEP-2000;	2000US-0230438;
PR	08-SEP-2000;	2000US-0231242;
PR	08-SEP-2000;	2000US-0231242;
PR	08-SEP-2000;	2000US-0231243;
PR	08-SEP-2000;	2000US-0231444;
PR	08-SEP-2000;	2000US-0231444;
PR	08-SEP-2000;	2000US-0231444;
PR	08-SEP-2000;	2000US-0231480;
PR	08-SEP-2000;	2000US-0232081;
PR	12-SEP-2000;	2000US-0232196;
PR	14-SEP-2000;	2000US-0232397;
PR	14-SEP-2000;	2000US-0232398;
PR	14-SEP-2000;	2000US-0232399;
PR	14-SEP-2000;	2000US-0232400;
PR	14-SEP-2000;	2000US-0232401;
PR	14-SEP-2000;	2000US-0233063;
PR	14-SEP-2000;	2000US-0233064;
PR	14-SEP-2000;	2000US-0233065;
PR	21-SEP-2000;	2000US-0234223;
PR	21-SEP-2000;	2000US-0234274;
PR	25-SEP-2000;	2000US-0234997;
PR	25-SEP-2000;	2000US-0234998;
PR	26-SEP-2000;	2000US-0234984;
PR	27-SEP-2000;	2000US-0235834;
PR	27-SEP-2000;	2000US-0235836;
PR	29-SEP-2000;	2000US-0236327;
PR	29-SEP-2000;	2000US-0236367;
PR	29-SEP-2000;	2000US-0236368;
PR	29-SEP-2000;	2000US-0236369;
PR	29-SEP-2000;	2000US-0236370;

PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249219.
PR	17-NOV-2000;	2000US-0249220.
PR	17-NOV-2000;	2000US-0249221.
PR	17-NOV-2000;	2000US-0249222.
PR	17-NOV-2000;	2000US-0249223.
PR	17-NOV-2000;	2000US-0249224.
PR	17-NOV-2000;	2000US-0249225.
PR	17-NOV-2000;	2000US-0249226.
PR	17-NOV-2000;	2000US-0249227.
PR	17-NOV-2000;	2000US-0249228.
PR	17-NOV-2000;	2000US-0249229.
PR	17-NOV-2000;	2000US-0249230.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251889.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.
XX		
PI	Rosen CA,	Barash SC, Ruben SW;
XX		
DR	WPI:	2001-465570/50.
XX	P-PDB;	AAM95378.
XT		

Isolated nucleic acid molecule encoding a reproductive system antigen

PT is used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 1; SEQ ID NO 134g; 1297bp + Sequence Listing; English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a coding sequence of the
 CC invention.
 XX
 SQ Sequence 632 BP; 190 A; 143 C; 166 G; 128 T; 5 other;

Query Match	56.5%	Score 429;	DB 22;	Length 632;
Best Local Similarity	89.4%;	Pred. No. 1.6e-124;		
Matches 505;	Conservative 0;	Mismatches 54;	Indels 6;	Gaps 4

0y 1 atgtgtagcaccacgtygggtgtgcacctgygaagaatccccctagtatgatgatgcccaaac 60
Db 52 atgtgtagcacccagtgggtgtgcacctggaagaatcccctagtatgatgatgccccaac 111

```

9y      accatagaattcaaatctcgcgctaatacacagacatgtcttcaaagaccacct 120
        |||
        |
Db      112 accatagaattcaaatctcgcgctaatacacagacatcatgtcttcaaagaccacct 171

```

Db 172 gctctctctcaccagaagctgctgagacaagaagtcttctccagagggcctgsgaat 231

232 tgttcagcaaatgcatgacagaggtgtcatgcttgcaagaattccaatccagt 291

292 gagaagggcataaaccttgagcaagaataagctctctcttggaagcattcttgagtagtg 351

361 gaatcccaagattgcacacacccacacacatattccaaatattcttataaaccttgaacacatcc 420
Db |||||
352 gagaaggaagatctcgacagacacgcctgcacaggtctctcgctcaaggtccaagagacttg 411
|||

Db 412 gaataccaagatcgcacag--cagcagltg--ccagtgcttctaacytgagcaagtg 468
 |||||
 Oy 421 ttgcagcatgaaacgtgtgagcccaaaatcatttcgattggccaaacgaatgactgaatt 480

Db 469 ctgagcatgaactgtgagcccaag-catttncatttgcaaccgagtagctgaatat 527
Oy 481 gttattccttgc--caccaccacaagcgaaccagcgagagcttcaagtccaaagaga 538

D6	528	ggtactccttgcgancaccacaagggaaccgagnaggaaggtttaaggccaagaag	587
O7	539	ttttgtaactaggggtctctcctt	563

Dd 588 attttggaactgangggcinttct 612

ABAA47100
ID ABAA47100 standard; DNA; 388 BP.
XX
AC ABAA47100:

AA	
DT	01-FEB-2002 (first entry)
XX	
DE	Human breast cell single exon nucleic acid probe #5795.

Human; microarray; single exon probe; gene expression; breast; disease; cancer; ss.

05 Homo sapiens.
xx

XX	MO200157271-A2.	XX
XX		XX
PD	09-AUG-2001.	PD
PF		PF
XX	30-JAN-2001; 2001WO-US00662.	XX
XX		XX
PR	04-FEB-2000; 2000US-0180312.	PR
PR	26-MAY-2000; 2000US-0207456.	PR
PR	30-JUN-2000; 2000US-0608408.	PR
PR	03-AUG-2000; 2000US-0632366.	PR
PR	21-SEP-2000; 2000US-0234687.	PR
PR	27-SEP-2000; 2000US-0236359.	PR
PR	04-OCT-2000; 2000GB-0024263.	PR
PA	(MOLE-) MOLECULAR DYNAMICS INC.	PA
PI		PI
PI	Penn SG, Hanzel DK, Chen W, Rank DR;	PI
XX		XX
DR	WPI; 2001-496933/54.	DR
XX		XX
PT	New spatially-addressable set of single exon nucleic acid probes,	PT
XX	useful for measuring gene expression in sample derived from human	XX
XX	breast, comprises number of single exon nucleic acid probes -	XX
XX	Claim 4; SEQ ID NO 5795; 327bp + sequence listing; English.	XX

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BR 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct-sequences.
 CC
 CC Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;
 CC
 CC

[illegible]

Db	Qy	Db
301	ggtcaaaaggaagcttgatataccaagattcgacacgccaagatggtccagtggtcttctt	360
406	aaacttgagcagctgcttggaacatgaag	433
361	aaacttgagcagctgcttggaacatgaag	388

RESULT	9
ABA64982	
ID	ABA64982 standard; DNA; 388 BP.

XX
AC
ABA64982;

01-FEB-2002 (first entry)

Human foetal liver single exon nucleic acid probe #13287.

Human; foetal liver; gene expression; single exon nucleic acid probe; ss

OS Homo sapiens.
XX

PN WO200157277-A2.
XX

PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US00669.
XX

PR	04-FEB-2000;	2000US-0180312.
PR	26-MAY-2000;	2000US-0207456.

PR	30-JUN-2000;	2000US-0608408.
PR	03-AUG-2000;	2000US-0632366.
PR	01-SEP-2000;	2000US-0634507.

PR	21-SEP-2000;	2000US-0234687
PR	27-SEP-2000;	2000US-0236359
PR	04-OCT-2000;	2000CB-0034363

04-001-2000; 2000GB-0024285.
PR
XX
BA (MOI E-) MOI ECHT AB DYNAMICS INC

PI Penn SG: Hanzel DK, Chen W, Rank DB:
FA (MODE) MOLECULAR DYNAMICS INC.
XX

WPI: 2001-483447/52

XX Human genome-derived single exon nucleic acid probes useful for PT

PT analyzing gene expression in human fetal liver -
XX

PS Claim 4; SEQ ID NO 13287; 639pp + sequence listing; English
XX

CC The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal

CC Liver The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human

CC probe of the invention.
CC fetal liver. The present sequence is a single exon nuclear acid
CC heterologous data for this patient did not form part of the

CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftn.int@wipo.int published not sequences
CC

Sequence 388 BP: 113 A: 91 C: 99 G: 85 T: 0 other:
CC 110M WIF0 ac 1cf: wifo: inc/ fad/ pub: 151nc01_p0c_000a0a
XX

0
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99

Query Match	51.1%;	Score 388;	DB 22;	Length 388;
Best Local Similarity	100.0%;	Pred. No. 1e-111;		

Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 gatgatvaccctaaccacatgaattcaaatcctcgctactacaccagacatcatgtc 105
|||||

Db 1 gatgatgcctaacaacatagaattcaaaatcctcgcctattacacccagacatcatgtc 60

QY 106 ttcaagagcaccctgctctctctcaccacaagcgcgtgagacaagaattgtccag 16

61 ttcaagagcaccctgcctctctctcaccacaagctgagacagaagtgtgtccag 120

QY 166 aggggcctggggaattgttcagcaaatgcatgacagaggtgtcattggtccctgcaga 22

Db 121 aggggcttgggaattgtctcagcaaatgltcatgagcagagtggtcatggtcttcaga 180
 QY 226 aatcccaatccagtggaagcagcataaactctggcaagaagaatctcttggaaagca 285
 Db 181 aatcccaatccagtggaagcagcataaactctggcaagaagaagaatctcttggaaagca 240
 QY 286 ttcttggagtagtggaaggaagatctgcagagcagcgtctgccaaggctctctgcag 345
 Db 241 ttcttggagtagtggaaggaagatctgcagagcagcgtctgccaaggctctctgcag 300
 QY 346 ggtcaaaagcgttggaatcccaagattccgacagcagcagtggtccaggtctcttct 405
 Db 301 ggtcaaaagcgttggaatcccaagattccgacagcagcagtggtccaggtctcttct 360
 QY 406 aacgtgagcagtgcttgagcatgaa 433
 Db 361 aacgtgagcagtgcttgagcatgaa 388
 RESULT 10
 ID ABA32090
 AC ABA32090 standard; DNA; 388 BP.
 AC ABA32090;
 DT 23-JAN-2002 (first entry)
 DE Probe #10556 for gene expression analysis in human heart cell sample.
 DE Human; gene expression; heart; microarray; vascular system; probe;
 DE cardiovascular disease; hypertension; cardiac arrhythmia;
 DE congenital heart disease; ss.
 OS Homo sapiens.
 OS
 XX WO200157274-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00666.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-488899/53.
 DR WPI; 2001-488899/53.
 PS Single exon nucleic acid probes for analyzing gene expression in human
 PS hearts -
 PS Claim 4; SEQ ID NO 10556; 530pp; English.
 The present invention relates to single exon nucleic acid probes for
 measuring human gene expression in a sample derived from human heart. The
 present sequence is one such probe. The probes may be used for
 predicting, measuring and displaying gene expression in samples derived
 from the human heart via microarrays. By measuring gene expression, the
 probes are useful for predicting, diagnosing, grading, staging,
 monitoring and prognosing diseases of the human heart and vascular system
 e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 congenital heart disease.
 Note: The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIFO
 at ftp.wifo.int/pub/published_pct_sequences.

XX SQ Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;
 Query Match 51.1%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. No. 1e-111;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 46 gatgtgaccttaaacacccatagatccaatctctgctactacaccagatcatgtc 105
 Db 1 gatgtgaccttaaacacccatagatccaatctctgctactacaccagatcatgtc 60
 QY 106 ttcaagagcaccctctctctctctcaccacaaagctgttggaacaaagttgtccag 165
 Db 61 ttcaagagcaccctctctctctctcaccacaaagctgttggaacaaagttgtccag 120
 QY 166 aggggcttgggaatgttccaagaatgagtcaltgacagaggtgtcatgtgcttcaga 225
 Db 121 aggggcttgggaatgttccaagaatgagtcaltgacagaggtgtcatgtgcttcaga 180
 QY 226 aatcccaatccagtggaagcagcataaactctggcaagaagaatcttctggaaagca 285
 Db 181 aatcccaatccagtggaagcagcataaactctggcaagaagaagaatcttctggaaagca 240
 QY 286 ttcttggagtagtggaaggaagatctgcagagcagcgtctgccaaggctctgcag 345
 Db 241 ttcttggagtagtggaaggaagatctgcagagcagcgtctgccaaggctctgcag 300
 QY 346 ggtcaaaagcgttggaatcccaagattccgacagcagcagtggtccaggtcttct 405
 Db 301 ggtcaaaagcgttggaatcccaagattccgacagcagcagtggtccaggtcttct 360
 QY 406 aacgtgagcagtgcttgagcatgaa 433
 Db 361 aacgtgagcagtgcttgagcatgaa 388
 RESULT 11
 ID AAK13408
 AC AAK13408 standard; DNA; 388 BP.
 AC AAK13408;
 DT 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe SEQ ID NO: 13399.
 DE Human; brain expressed exon; gene expression analysis; probe;
 DE microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 DE epilepsy; cancer; ss.
 OS Homo sapiens.
 OS
 XX WO200157275-A2.
 XX PN 09-AUG-2001.
 XX PD 30-JAN-2001; 2001WO-US00667.
 XX PF 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-483446/52.
 DR WPI; 2001-483446/52.

PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 13399; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.

XX Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;

Query Match 51.1%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. No. 1e-111;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 46 gatgatgacttaaacacataagaattcaaatctctgctactacacagacatcatgtc 105
 DB 1 gatgatgacttaaacacataagaattcaaatctctgctactacacagacatcatgtc 60
 XX 106 ttcaagagacccctgctctctctccacaaagctgtgagacaagaattgtccag 165
 DB 61 ttcaagagacccctgctctctctccacaaagctgtgagacaagaattgtccag 120
 XX 166 agggcctgagggaattgttcacaaatgacatgacagaggtgtcatgcttcagca 225
 DB 121 agggcctgagggaattgttcacaaatgacatgacagaggtgtcatgcttcagca 180
 XX 226 aattcccaatccagtgagagagcacaataaccttggcaagaagaattctcttggaaaga 285
 DB 161 aattcccaatccagtgagagagcacaataaccttggcaagaagaattctcttggaaaga 240
 XX 286 ttctttgagtagtgagagagagcagcagccttgcacaaagtctctgtccag 345
 DB 241 ttctttgagtagtgagagagagcagcagccttgcacaaagtctctgtccag 300
 XX 346 ggtcaaaagacgttggagaaatcccaagatcgcacagcagcagtggttccagtgcttct 405
 DB 301 ggtcaaaagacgttggagaaatcccaagatcgcacagcagcagtggttccagtgcttct 360
 XX 406 aacgttgagcagtgcttggagcagtgaaag 433
 DB 361 aacgttgagcagtgcttggagcagtgaaag 388

XX T 12
 XX 145
 XX ID AAK39145 standard; DNA; 388 BP.

XX AAK39145;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 13702.

XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukemia; lymphoma; myeloma; ss.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024253.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 13702; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention.

XX Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;

Query Match 51.1%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. No. 1e-111;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 46 gatgatgacttaaacacataagaattcaaatctctgctactacacagacatcatgtc 105
 DB 1 gatgatgacttaaacacataagaattcaaatctctgctactacacagacatcatgtc 60
 XX 106 ttcaagagacccctgctctctctccacaaagctgtgagacaagaattgtccag 165
 DB 61 ttcaagagacccctgctctctctccacaaagctgtgagacaagaattgtccag 120
 XX 166 agggcctgagggaattgttcacaaatgacatgacagaggtgtcatgcttcagca 225
 DB 121 agggcctgagggaattgttcacaaatgacatgacagaggtgtcatgcttcagca 180
 XX 226 aattcccaatccagtgagagagcacaataaccttggcaagaagaattctcttggaaaga 285
 DB 181 aattcccaatccagtgagagagcacaataaccttggcaagaagaattctcttggaaaga 240
 XX 286 ttctttgagtagtgagagagagcagcagccttgcacaaagtctctgtccag 345
 DB 241 ttctttgagtagtgagagagagcagcagccttgcacaaagtctctgtccag 300
 XX 346 ggtcaaaagacgttggagaaatcccaagatcgcacagcagcagtggttccagtgcttct 405
 DB 301 ggtcaaaagacgttggagaaatcccaagatcgcacagcagcagtggttccagtgcttct 360
 XX 406 aacgttgagcagtgcttggagcagtgaaag 433
 DB 361 aacgttgagcagtgcttggagcagtgaaag 388

RESULT 13
 AAI19954
 ID AAI19954 standard; DNA; 388 BP.

XX AAI19954;

XX 12-OCT-2001 (first entry)

XX Probe #9887 for gene expression analysis in human cervical cell sample.

XX Probe: human; microarray; gene expression; cervical epithelial cell;

XX cervical cancer; ss.

XX Homo sapiens.

XX XX WO200157278-A2.
 XX XX 09-AUG-2001.
 XX XX 30-JAN-2001; 2001WO-US00670.
 XX XX 04-FEB-2000; 2000US-0180312.
 XX XX 26-MAY-2000; 2000US-0207456.
 XX XX 30-JUN-2000; 2000US-0608408.
 XX XX 03-AUG-2000; 2000US-0632366.
 XX XX 21-SEP-2000; 2000US-0234687.
 XX XX 27-SEP-2000; 2000US-0236359.
 XX XX 04-OCT-2000; 2000GB-0024263.
 (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PT WPI; 2001-488901/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX Claim 25; SEQ ID NO 9887; 487pp; English.
 XX The present invention relates to human single exon nucleic acid probes
 CC (SENP). The present sequence is one such probe. The SENPs are derived
 CC from human Hela cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging
 CC of diseases of the cervix, notably cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;
 SQ

Query Match 51.1%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. NO. 1e-111;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 gatgatgacctaacaacacatgaattcaaatccctcgctactacacagacatcatg 105
 D 1 gatgatgacctaacaacacatgaattcaaatccctcgctactacacagacatcatg 60
 QY 106 ttcaagagcaccctgctctctctcaccagaagctgctgagaacaagaattgtccag 165
 D 61 ttcaagagcaccctgctctctctcaccagaagctgctgagaacaagaattgtccag 120
 QY 166 aggggctgggaattgttcgaagaatgacatgacagagtggtcattgaccttcaga 225
 D 121 aggggctgggaattgttcgaagaatgacatgacagagtggtcattgaccttcaga 180
 QY 226 aattcccaatccagttgagaagcacaataaccttgcaagaagaattctcttggaaaga 285
 D 181 aattcccaatccagttgagaagcacaataaccttgcaagaagaattctcttggaaaga 240
 QY 286 ttcttggagtagtgagaagaagatcgacagacagctgccaaggtctctgctcag 345
 D 241 ttcttggagtagtgagaagaagatcgacagacagctgccaaggtctctgctcag 300
 QY 346 ggtcacaagagcgttggagaattcgacaagcagccagcagtggtccaggtcttctc 405
 D 301 ggtcacaagagcgttggagaattcgacaagcagccagcagtggtccaggtcttctc 360
 QY 406 aacgtgagcagtgcttggagcagatgaag 433
 D 361 aacgtgagcagtgcttggagcagatgaag 388

RESULT 14
 AA145152
 ID AA145152 standard; DNA; 388 BP.
 XX AC AA145152;
 XX DT 17-OCT-2001 (first entry)
 XX DE Probe #13838 used to measure gene expression in human placenta sample.
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX KW genetic disorder; ss.
 XX OS Homo sapiens.
 XX EN WO200157272-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00663.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PT WPI; 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX Claim 25; SEQ ID NO 13838; 654pp; English.
 XX The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;
 SQ

Query Match 51.1%; Score 388; DB 22; Length 388;
 Best Local Similarity 100.0%; Pred. NO. 1e-111;
 Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 gatgatgacctaacaacacatgaattcaaatccctcgctactacacagacatcatg 105
 D 1 gatgatgacctaacaacacatgaattcaaatccctcgctactacacagacatcatg 60
 QY 106 ttcaagagcaccctgctctctctcaccagaagctgctgagaacaagaattgtccag 165
 D 61 ttcaagagcaccctgctctctctcaccagaagctgctgagaacaagaattgtccag 120
 QY 166 aggggctgggaattgttcgaagaatgacatgacagagtggtcattgaccttcaga 225
 D 121 aggggctgggaattgttcgaagaatgacatgacagagtggtcattgaccttcaga 180
 QY 226 aattcccaatccagttgagaagcacaataaccttgcaagaagaattctcttggaaaga 285
 D 181 aattcccaatccagttgagaagcacaataaccttgcaagaagaattctcttggaaaga 240
 QY 286 ttcttggagtagtgagaagaagatcgacagacagctgccaaggtctctgctcag 345
 D 241 ttcttggagtagtgagaagaagatcgacagacagctgccaaggtctctgctcag 300

QY 346 ggtcaaaagacgttggaataccaagattcgacagccagctggttccagggtcttct 405
|||||
DB 301 ggtcaaaagacgttggaataccaagattcgacagccagctggttccagggtcttct 360
|||||
QY 406 aacgttgagacagctgcttgagacatgaag 433
|||||
DB 361 aacgttgagacagctgcttgagacatgaag 388
|||||

RESULT 15
AAI05666 standard; DNA; 388 BP.
XX AAI05666;
XX AC AAI05666;
XX DT 09-OCT-2001 (first entry)
XX DE Probe #5657 used to measure gene expression in human breast sample.
XX KE Probe; human; breast disease; breast cancer; development disorder; ss;
XX KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX O Homo sapiens.
XX PN WO200157270-A2.
XX PD 09-AUG-2001.
XX PF 29-JAN-2001; 2001MO-US00661.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-476286/51.
XX PT Novel single exon nucleic acid probe used to measuring gene expression
XX PR in a human breast -
XX PS Claim 25; SEQ ID NO 5657; 322pp; English.
XX SQ

CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pft_sequences.
XX SQ

Sequence 388 BP; 113 A; 91 C; 99 G; 85 T; 0 other;

Query Match 51.1%; Score 388; DB 22; Length 388;
Best Local Similarity 100.0%; Pred. No. 1e-111;
Matches 388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 gatgatgacttaaacaccatagaattcaaatcctcgctactaataccagacatcatgtc 105
|||||
DB 1 gatgatgacttaaacaccatagaattcaaatcctcgctactaataccagacatcatgtc 60
|||||

QY 106 ttcaagaacacccctgctctctctctcaccacaagctgtgagaacaagaattgtccag 165
|||||
DB 61 ttcaagaacacccctgctctctctctcaccacaagctgtgagaacaagaattgtccag 120
|||||
QY 166 aggggcttgaggaaattgttcagaacaatgagtcataggacagaagtgcccttcaga 225
|||||
DB 121 aggggcttgaggaaattgttcagaacaatgagtcataggacagaagtgcccttcaga 180
|||||
QY 226 aattcccaatccagtgagaaaggccataaaccttggcaaaaagcttcttgaaagca 285
|||||
DB 181 aattcccaatccagtgagaaaggccataaaccttggcaaaaagcttcttgaaagca 240
|||||
QY 286 ttcttggagtagtgagaaaggcattcgcagagcagccttgccaaggtctctgcag 345
|||||
DB 241 ttcttggagtagtgagaaaggcattcgcagagcagccttgccaaggtctctgcag 300
|||||
QY 346 ggtcaaaagacgttggaataccaagattcgacagccagcagctggttccagggtcttct 405
|||||
DB 301 ggtcaaaagacgttggaataccaagattcgacagccagcagctggttccagggtcttct 360
|||||
QY 406 aacgttgagacagctgcttgagacatgaag 433
|||||
DB 361 aacgttgagacagctgcttgagacatgaag 388
|||||

Search completed: June 20, 2002, 05:05:37
Job time: 5981 sec

VERSION	AX174785.1	GI:14598301
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
REFERENCE	1 (bases 1 to 954)	
AUTHORS	Reed, J.C. and Godzik, A.	
TITLE	Bcl-2 polydeptides: encoding nucleic acids and methods of use	
JOURNAL	Patent: WO 0144282-A 3 21-Jun-2001; The Burnham Institute (US)	
FEATURES	Location/Qualifiers	
source	1..954	
CDS	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	196..954	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="C4C43801.1"	
	/db_xref="GI:14598302"	
	/translation="MCSGSCDLEIEIPDDDLNIEKILAYVRHNFVKSYPALFS PLILRSLRSORRIGLNCNANESWTEVSPICNNSOSEKAINGKRSKMAFVVEK EBSOSPPAKYASAGORTLEXYODSHSOOMRSLNVEOCEHFAVDKVISIANRVAEI VSWMPPOATVAGGFSKLEIFVEFGSLSPOLGHHVYVASSKDEDEQILAKIVELKY SSDQLERKRTAFPIPIPIVDTSISGFPPDGLMACI"	
BASE COUNT	267 a 237 c 240 g 210 t	
ORIGIN		
Query Match	99.8%	Score 757.4; DB 6; Length 954;
Best Local Similarity	99.98%	Pred. No. 3.6e-220;
Matches 758; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY 1 atgtgtgacacccgctggtgtgtgacctggaagaatcccccctagatgatgacctaac 60		
DB 196 AAGTGTAGACCCAGTGGGTGTGACCTGGAAGAAATCCCCCTGATGATGATGACCTAAC 255		
QY 61 accatgaattcaaatctctcgctactacaccagacatcatgtcttcaagacacct 120		
DB 256 ACCATGGAATTCAAAATCCCGCTCTACACACCAACATCATGTCTTCAAGACACCCCT 315		
QY 121 gctctcttcacacaagctgtctgtagaacaagaattgttccagaagggtcctgggaat 180		
DB 316 GCTCTCTTCTCACCAAGCTGCTGAGAACAGAAATTTGTCACAGAGGGGCTGGGGAAAT 375		
QY 181 tgttcagcaaatgagtcattgtagacagagggtgtcatgtgaccttcgagaattcccaatcagt 240		
DB 376 TGTTCAGCAAAATAGTGTATGTGACAGAGGTGTGATGCGCTTGCAGAAATTCACATCCAGT 435		
QY 241 gagaaggcctaataaccttggcaagaanaagctcttcttggaagcattcttggagtagtg 300		
DB 436 GAGAAAGCCATTAACCTCTGGCCAAAGAAAAGTCTTTGGAAAGCATTTTGGAGTAGTG 495		
QY 301 gagaagaagatctcgagagcagcagcctgccaaggtctctgtccaggtgtaagaagcgttg 360		
DB 496 GAGAAAGAAAGATTGCGACAGCACCGCTGCCAAGGTCCTGTCTCAGGGTCAAAAGACGTTG 555		
QY 361 gaatacacaagatctgcacagccagcagcagtggtgtccaaagtcttcttaagtggaagcagtg 420		
DB 556 GAATACCAAGATTTCGCACAGCACAGTGTGTCCAGGTGCTTTCTTAACGTGGAGCAATGCC 615		
QY 421 ttggagcatgaagcgtgtgaccaccaagtcattctcatgtgcaaccgagtagtgaat 480		
DB 616 TTGGACATGAACTGTGGAGCCCAAAAGTCAATTTCCATTGCAACCGAGTAGCTGAATTT 675		
QY 481 gtttattctccgtggcaccacacacaagaaggacccaaggaaggcttcaagtccaagagatt 540		
DB 676 GTTTACTCTCGGCACACCAACAAGCGCCCAAGGAGGAGGCTTCAAGTCCAAAGAAATTT 735		
QY 541 ttgttaactgaagggtctctcccttcacagtcacaaggccagtcgtctgtagcttaagctct 600		
DB 736 TTTGTAACTGAGGGGTCTCTCTTCCAGTCCAAAGCCCAAGTGCCTGTACCTTAAGTTCT 795		

QY	601	aaagaagaatgaagaagaacaactactagacaaatggttgagctgtgcgaatattcaga	660
Db	796	AACGAAAGATGAGAGAGACAACTACTAGCCAAATGTTGAGCTGCTGAATATTACAGA	855
QY	661	gatacagcttggaagaagaagacactgccttcataccatcccttggttgacacagcattc	720
Db	856	GATCAGCTTGGAAGAAAGACACTGCTTCATCCCATTCCTGTTGACACAGCATC	915
QY	721	cagggtttccacagatgattgatgagctgcatttga	759
Db	916	CAGGGTTTCCACAGAGATGTTGATGGCTGCATTTGA	954
RESULT	4		
AF281255		AF281255	2039 bp mRNA linear PRI 28-JAN-2001
LOCUS			
DEFINITION			Homo sapiens apoptosis regulator BCL-G short form (BCLG) mRNA,
ACCESSION			complete cds.
VERSION			AF281255
KEYWORDS			AF281255.1 GI:12584086
SOURCE			human.
ORGANISM			Homo sapiens
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS			1 (bases 1 to 2039)
TITLE			Guo,B., Godzik,A. and Reed,J.C.
JOURNAL			Bcl-2, a novel pro-apoptotic member of the bcl-2 family
MEDLINE			J. Biol. Chem. 276 (4), 2780-2785 (2001)
REFERENCE			2 (bases 1 to 2039)
AUTHORS			Guo,B., Godzik,A. and Reed,J.C.
TITLE			Direct Submission
JOURNAL			Submitted (22-JUN-2000) The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES			
source			Location/Qualifiers
			1..2039
			/organism="Homo sapiens"
			/db_xref="taxon:9606"
			/chromosome="12"
			1..2039
			/gene="BCLG"
			196..954
			/gene="BCLG"
			/note="member of Bcl-2 family"
			/codon_start=1
			/product="apoptosis regulator BCL-G short form"
			/protein_id="AAG59794.1"
			/db_xref="GI:12584087"
			/translation="MCSSTSGDLEIPDDDLNTEFKILAYTRHHVFKSTPALEFS PKILRTLSQRLGNCSSANSWEVSPNRSOSESEKAINLGRKKSWMKAFGVVEK EDSMPSPKAYSAGORFLEYODSHSOQSRCSNVEQCLEHEAVDPKVISIANRVAEII VVSPSPQATQAGGSKSEIFETVETSLSPOLGSHVAVASSKDEEDQILAKIYELLKYI SGQLEKRDIAFIPIPLVDYISIOGPPDGLMACT"
BASE COUNT	567 a	475 c	484 g 513 t
ORIGIN			
Query Match	99.8%	Score 757.4	DB 9: Length 2039;
Best Local Similarity	99.9%	Pred. No. 3.9e-220;	
Matches 758; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1	atbtgttagacacacatggtgtgtcactgtgaagaatccccctagatgatgatgacctaac	60
Db	196	ATGTGTGACACACCAATCGGTGTGACCTGGAAGAANAATCCCTAGATGATGACCTAAAC	255
QY	61	accatagaatccaatcctgcctactacacacagacatcatgcttccaagaagaccct	120
Db	256	ACCATAGAATTCAAAAATCCCTGCGCTACTACACACACATCATGTCTTCAAGACACACCCCT	315
QY	121	gtctctcttcacacaaagctgtctgagaacaaagaattgttccagaggggctctggggaat	180
Db	316	GCTCTCTTCTCACCACAAAGCTCTGTGAGAACAGAGATTGTCTCCAGAGGGGCTGGGGAAT	375

QY 181 ttctcagaatgaatgcatgacagaaggtgtcattgagccttgacagaattcccaatccact 240
 |||||||
 Db 376 TGTTCAGCAATGATGATGACAGAGGTGTCAAGGCTTGAGAAATTCACATCCAGT 435
 |||||||
 QY 241 gagaagccaataaccttggcaagaagaagctctcttgaaagcctctcttgagtagtg 300
 |||||||
 Db 436 GAGAAAGCCATTAACCTTGGCAAGAAAGTCTCTTGAAAGCATTTCTTGAGTAGTG 495
 |||||||
 QY 301 gagaagaagaatctgcagagacagccttgcaaggtctctgctcaaggttcaagaagagcttg 360
 |||||||
 Db 496 GAAAGGAAATTCGCGAGACACGCTGCCAAGGTCTCTGCTCAGAGGTCAAGAGAGCTTG 555
 |||||||
 QY 361 gaatacaagattcgcagcagcagcagcagtggtgtccttctcaagcttgagcagtg 420
 |||||||
 Db 556 GAATACCAAGATTGCGACAGCCAGCAGTGTGTCAAGTCTTTCTTACGTGGAGCAGTGC 615
 |||||||
 QY 421 ttgagcagatgaagctgtgagaccccaagtcattccattgccaacagtagtgaatt 480
 |||||||
 Db 616 TTGAGCATGAGCTGTGAGACCCCAAGTCAATTCATTGCCAACCGAGTAGCTGAATTT 675
 |||||||
 QY 481 gttattctctgcccacccaagaagcagcagcagcagcagcagcagcagcagcagcagc 540
 |||||||
 Db 676 GTTTACTCTCGCCACACCAACAAGCCAGCGAGGAGGCTTCAAGTCCAAAGAGATT 735
 |||||||
 QY 541 ttgttaactgaaggtctctctctcctcagctccaagcagcagcagcagcagcagcagc 600
 |||||||
 Db 736 TTTGTACTGAGGGTCTCTCTCTTCCAGCTCCAAAGGCCACGTGCTGTACTTCAAGTTCT 795
 |||||||
 QY 601 aagaagaatgaagaagaacaataactagccaagaattgttgagctgtgaaatatacaga 660
 |||||||
 Db 796 AAGAAAGATGAAGAAGAAACAATGACTAGCCAAATTTGTTGACTGCTGAAATATTTCAGA 855
 |||||||
 QY 661 gatcagttgaaagaagaagacacgtccttcatcccatctcccttggttgacaccagatc 720
 |||||||
 Db 856 GATCAGTTGGAAGAAAGAGACGCTTCATCCCATTTCCCTGGTGAACACGACATC 915
 |||||||
 QY 721 caaggtttccacagagatgttgatgagcctgattga 759
 |||||||
 Db 916 CAGGCTTTCCACAGATGCTTGTGATGCTGCAATTGA 954
 |||||||

RESULT 5
 AY040274 1176 bp mRNA linear PRI 01-AUG-2001
 LOCUS Homo sapiens apoptosis regulator BCL-2 median form (BCL2) mRNA,
 DEFINITION complete cds.
 ACCESSION AY040274 GI:15072494
 VERSION AY040274.1
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1176)
 AUTHORS Montpetit, A., Bolly, G. and Sinnott, D.
 TITLE A detailed transcriptional map of the chromosome 12p12 tumor
 suppressor locus
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1176)
 AUTHORS Montpetit, A. and Sinnott, D.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUN-2001) Hemato-Oncology, Hopital Ste-Justine, 3175
 Cote-Ste-Catherine, Montreal, QC H3T 1C5, Canada
 FEATURES
 source
 1.1176
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="12"
 /map="12p12"
 1.1176
 /gene="BCL2"
 8.838

gene
CDS

/gene="BCL2"
 /note="member of the BCL-2 family"
 /codon_start=1
 /product="apoptosis regulator BCL-2 median form"
 /protein_id="AAK72109.1"
 /db_xref="GI:15072495"
 /translation="MCSISGCDLEIEIPDDDLNLTIEFKILAYTRHHVKSYPALFS
 PKLIRTRLSQRGNGNSANESWVNCRSOSSEKALINLGRKSKSVKAFGVYK
 EDSSTPAKVSAQGRRLTYODSHSDQMSRCLSNVQCLHEHVDKYSIANRVAEI
 VYSPPPOATAGCKRKSKEIYTBELSLDLOSHVPAASSKSKDEEQILAKIYELLKI
 SGDLEKREKWEESAKTKIKMALLMRSSGGSPRVSLTPEAFKNSFTLTPGFSHN"
 BASE COUNT 363 a 277 c 270 g 266 t
 ORIGIN

Query Match 89.4%; Score 678.8; DB 9; Length 1176;
 Best Local Similarity 99.7%; Pred. No. 3.9e-196;
 Matches 680; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgtgtagcaccagtggtgtgtgacctggaagaatcccccataatgtagtaccctaac 60
 |||||||
 Db 8 ATGTGTACACACAGTGGGTGTGACCTGGAAGAATCCCTAGATGATGATGACTTAAC 67
 |||||||
 QY 61 accatagaattcaaaatccctgcctactacacccaagacatgcttccaagaagaccct 120
 |||||||
 Db 68 ACCATGAGATTCAAAATCTCGCTACTACTACACGACATCATCTTTCAGAGACCCCT 127
 |||||||
 QY 121 gctctcttcccaagaagctgtgagaacaagaagttgtcccaagaaggccttgaggat 180
 |||||||
 Db 128 GCTCTCTTCTACCAAGAGCTGTGAGAACAGAAAGTTGTGCCAGAGGGGCTGGGGAA 187
 |||||||
 QY 181 ttgtcagaagaatgagtcattgagaagaaggtgtcattgagccttgagaagaattcccaatcag 240
 |||||||
 Db 188 TGTTCAGCAATGATGATGACAGAGGTGTCTAGGCTCTTGCAAAATTCCTCAATCCAGT 247
 |||||||
 QY 241 gagaagccaataaccttggcaagaagaagctctcttgaaagcattcttggagtagtg 300
 |||||||
 Db 248 GAGAAAGCCATTAACCTTGCGCAAGAAAGTCTTCTTGAAAGCATCTTGTGAGTAGTG 307
 |||||||
 QY 301 gagaagaagaatctgcagagacagcctgcacaaggtctctgtcagaggttcaagaagcgttg 360
 |||||||
 Db 308 GAGAAGGAAGATTCCAGAGACGACGCTGCCAAGGTCTGTGCTCAAGGCTCAAGAGCAGTTG 367
 |||||||
 QY 361 gaatacaagattcgcagcagcagcagcagtggttccaggtgtcttctcaacgtggagcagtg 420
 |||||||
 Db 368 GAATACCAAGATTCCAGACGACGACGAGTGTCCAGGTCTTTCTTAACGTGGAGCAGTGC 427
 |||||||
 QY 421 ttgagcagatgaagctgtgagccccaagtcattccattgccaacagtagtgaatt 480
 |||||||
 Db 428 TTGAGCATGAGGCTGTGTGAGCCCAAGATGATTCATTCCTCAACCGAGTAGTGAATTT 487
 |||||||
 QY 481 gttattctctgcccacccaagaagcagcagcagcagcagcagcagcagcagcagcagc 540
 |||||||
 Db 488 GTTTACTCTCGCCACACCAAGGAGCCAGGAGGAGGAGCTTCAAGTCCAAAGAGATT 547
 |||||||
 QY 541 ttgttaactgaaggtctctctctcagcctccaagcagcagcagcagcagcagcagcagc 600
 |||||||
 Db 548 TTTGTAAGTGAAGGTCTCTCTTCCAGCTCCAGGACGACGCTGTAGCTTCAAGTCTT 607
 |||||||
 QY 601 aagaagaatgaagaagaacaataactagccaagaattgttgagctgtcgaataatcaga 660
 |||||||
 Db 608 AAGAAAGATGAAGAAGAAACAATGACTAGCCAAATTTGTTGAGCTGTGAATATATTCAGA 667
 |||||||
 QY 661 gatcagttggaagaagaagaca 682
 |||||||
 Db 668 GATCAGTTGGAAGAAGAGAAA 689
 |||||||

RESULT 6
 AX207692 984 bp DNA linear PAT 31-AUG-2001
 LOCUS AX207692
 DEFINITION Sequence 1 from Patent WO0157213.

ACCESSION AX207692
VERSION AX207692.1 GI:15422374
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
REFERENCE 1 (bases 1 to 984)
AUTHORS Donoho, G., Hilbun, E., Turner, C.A., Friedrich, G.B., Abuin, A.,
Zambrowicz, B., and Sands, A.T.
TITLE Human bcl-x-like proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0157213-A 1 09-AUG-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source Location/Qualifiers
1..984
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 289 a 230 c 250 g 215 t
ORIGIN

Query Match 89.3%; Score 678; DB 6; Length 984;
Best Local Similarity 100.0%; Pred. No. 6,7e-196; Mismatches 0; Indels 0; Gaps 0;

Matches 678; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgtgtgacacagtggtgtgtgaccttggaagaatccccctagatgatgtacctaac 60
DB 1 ATGTGTGACACAGTGGGTGTGACTGGAGAAATCCCCCTAGATGATGACTTAAC 60
QY 61 accatagaatcaaatcctctgctactacacacacacacacacacacacacaccc 120
DB 61 ACATATAAATTTCAAAATCCTGCTGCTACACACACACATCATGTCTTCAAGACACCCCT 120
QY 121 gctctctctccacaaagctctggaagaagtttctccagaagggcctggaagat 180
DB 121 GCTCTCTTCTCCAAAGCTCTGGAACAGAAAGTTTCTCCAGAGGGCCTGGGGAAT 180
QY 181 tttcagaacaatgacatgacagaggtgtcatggtccttgagaatccccatcagt 240
DB 181 TGTTCACCAATGAGTATGACAGAGGTGTCTATGCGCTTGCAGAAATCCCAATCCAGT 240
QY 241 gagaagagcctaacaacctgtgcaagaagaagttctcttgagaagcattcttgagtagtg 300
DB 241 GAGAAGGCCATTAACCTGTGGCAAGAAAAAGTCTTGTGAAAGCATTTCTTGGAGTAGTG 300
QY 301 gagaagaagatctgcagagacgcctgcacaaagctctgtcagaggttcaagaagcgttg 360
DB 301 GAGAAGGAATTTGGCAGACAGACGCTGCCAAGTCTCTGCTCAGGGTCAAGAGGAGCTTG 360
QY 361 gaatccaagatctgcacagcagcagtggtctcaggtgtcttcttaacgttgagcagtg 420
DB 361 GAATACCAAGATTTGCACAGCAGAGAGTGTCTCAAGTCTTCAACGAGAGTGTG 420
QY 421 ttgagagatgaagtggtgagaccccaaatcattcattcagcacaacgcagtagtgaatt 480
DB 421 TTGAGAGATGAAGTGTGGACCCCAAAAGTCAATTTCCATTGCAACCGAGTAGTGAATTT 480
QY 481 gtattatccctgagcacaac 540
DB 481 GTTATTCTCGGCCAC 540
QY 541 ttgttaactagaggttctctcctcctcagcttcaaggtcagtgctgttaagcttcaagttc 600
DB 541 TTGTAACTGAGGGTCTCTCTTCCAGCTCCAAAGGCCAGTGTCTTACTTCAAGTCT 600
QY 601 aagaagaatgaagaagaacaatactagcacaatctgttgagctgtgaataattcaaga 660
DB 601 AAGAAAGATGAAGAAGAACAAATCTAGCCAAATTTGTTGAGCTGTGAATAATTTCAGGA 660
QY 661 gatcagcttggaagaagaag 678
DB 661 GATCAGTTGGAGGAAG 678

RESULT 7
LOCUS AX174783 1179 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 1 from Patent WO0144282.
ACCESSION AX174783
VERSION AX174783.1 GI:14598299
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
REFERENCE 1 (bases 1 to 1179)
AUTHORS Reed, J.C. and Godzik, A.
TITLE Bcl-2 polypeptides, encoding nucleic acids and methods of use
JOURNAL Patent: WO 0144282-A 1 21-JUN-2001;
The Burnham Institute (US)
FEATURES
Source Location/Qualifiers
1..1179
/organism="Homo sapiens"
/db_xref="taxon:9606"
196..1179
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC43798.1"
/db_xref="GI:14598300"
/translation="MSTSGDLEETPIDDDIINTFEKILAYVTRHVEKSPALFS
PKLITRSLISQGLGCSANESWTEVSMPCRNOSSEKAIINLKKSSWAFVEYK
EDSSTPAKVSAGQRTLEYODSHSOOWSCULSNVECDLEADVPVSTIANVLEI
VYSWPPQATGAGFRSEKIEIYVEKGFSEFQDGHVPAVSSSKKDEEDILAVILELKY
SGDLEKRLKDKRALMGHFOODGLSYSEVFKITDQVLMGVPRGSEVKAQGFKAALVI
DVTATLTALDHNPMNIVLFGFTYLEKFNFSFWIQQHGWKRIIGISHREVQ"

CDS

Query Match 89.1%; Score 676.4; DB 6; Length 1179;
Best Local Similarity 99.9%; Pred. No. 2.1e-195;
Matches 677; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atgtgtgacacagtggtgtgtgaccttggaagaatccccctagatgatgtacctaac 60
DB 196 ATGTGTGACACAGTGGGTGTGACTGGAGAAATCCCCCTAGATGATGACTTAAC 255
QY 61 accatagaatcaaatcctctgctactacacacacacacacacacacacacaccc 120
DB 256 ACCATAGAATTTCAAAATCCTGCTCTACTACACACACATCATGTCTTCAAGAGCACCCCT 315
QY 121 gctctctctccacaaagctgtgtagaagaagttgttccagaagggcctggaagat 180
DB 316 GCTCTCTTCTCAACCAAGGTCGTGAGAACAAAGTTTGTCTCCAGAGGGCCTGGGGAAT 375
QY 181 tttcagaacaatgacatgacagaggtgtcatggtccttgagaatccccatcagt 240
DB 376 TGTTCAGCAAAAGATGACATGACAGAGGTGTCAATGGCTTGCACAAATTTCCCAATCCAGT 435
QY 241 gagaagggcctaacaaccttgcaagaagaagttctcttggaagaagcatttgagtagtg 300
DB 436 GAGAAGGCCATTAACCTTGGCAGAAAGAAAGTCTTGTGAAAGCATTTCTTGGAGTAGTG 495
QY 496 gagaagaagatctgcagagcagcctgccaaggtctctcagaggttcaagaagcgttg 360
DB 496 GAGAAGGAATTTCCGAGAGCAGCCTGCCAAGGTCTCTGCTCAGGGTCAAGAGAGCTTG 555
QY 361 gaatccaagatctgcacagcagcagtggttccaggtgtcttcttaacgttgagcagtg 420
DB 556 GAATACCAAGATTTCCGACAGCAGCAGTGTCTCAGGTCTTCTTAAGCTGAGACAGTGC 615
QY 421 ttgagagatgaagctgtggaaccccaagatcattcattcagcacaacgcagtagtgaatt 480
DB 616 TTGAGCATGAAGCTGTGGACCCCAAGTCAATTTCCATTGCAACCGAGTAGTGAATTT 675
QY 481 gtattatccctgagcacaac 540

Db	676	GTTTACTCCCTGGCACCACCAAGGCGAGGCCGAGGGCTTCGAATCCAAAGAGATT	735
Oy	541	tttgttaactgagggctctcctccctcaagtcgaaggccaagtgcctgttagcttcaagtctc	600
Db	736	TWTGTACTGAGGGCTCTCTCTCCAGCTCCCAAGGCACAGTGCTGTAGCTCAAGTTC	795
Oy	601	aagaatcatgaagaacacaatactagccaaaattgttgagctcgctgaatatlcagga	660
Db	796	AAGAAGATCATACAAGAACAAATACTACCACAAAATTGTTGAGCTGCTGAATATTTCAGGA	855
Oy	661	gatcaagttggaagaagaag	678
Db	856	GATCAGCTTGGAAGAAAG	873

RESULT 8

AF281254 LOCUS AF281254 1886 bp mRNA linear PRI 28-JAN-2001

DEFINITION Homo sapiens apoptosis regulator BCL-G long form (BCLG) mRNA,

complete cds.

AF281254 AF281254 GI:12584084

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homimidae; Homo.

REFERENCE 1 (bases 1 to 1886)

Guo,B., Godzik,A. and Reed,J.C. Bcl-g, a novel proapoptotic member of the bcl-2 family J. Biol. Chem. 276 (4), 2780-2785 (2001)

MEDLINE 21264734

2 (bases 1 to 1886)

Guo,B., Godzik,A. and Reed,J.C. Direct Submission Submitted (22-JUN-2000) The Burnham Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

JOURNAL LOCATION/Qualifiers

FEATURES source

1..1886 /organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="12"

1..1886 /gene="BCLG"

/gene="BCLG"

196..1179 /note="member of Bcl-2 family"

/codon_start=1 /product="apoptosis regulator BCL-G long form"

/protein_id="AAG59793.1"

/db_xref="GI:12584085"

/translation="WCSTSGCDLEIPDDDLNTIEKILAYTRHHVKSYPALFS PKLRITSLSRIGNCNSANSWTEVSMPCRNSQSEKAINTGKKSSMAFEGVVERK EDSOSPAPKVASOAGORTLEYODSHSQMRCISNVEOCIEHEHADRPVISIANVAEIDH VYSWPPOATAGAGFTKEIYEVTBEGLSPOLGHVPVAVASSSKKEEDD ILAKIVLLKYK SGOLEERKLKKDKRALMGHPDGLSYSVKRTINDOYLGMGDPRRESEVKAGGPKALVI DYTAKFAIDNNPNRNVLGFEFTKTKLENKPSFMIOQHGNKILGISHEVD"

BASE COUNT 535 a 455 c 450 g 466 t

ORIGIN

Query Match	89.1%	Score 676.4:	DB 9:	Length 1886:
Best Local Similarity	99.9%:	Pred. No. 2.2e-195;		
Matches 677;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

Oy	1	atgtgtagaccacgctgggtgtgacctgggaagaatccccctagatgatgataaac	60
Db	196	ATGTGTAGACCAACGATGGGTGTGACTCGAAGAAATCCCCCTAGATGATGACTTAAC	255
Oy	61	accatgaatcaaatctcgtcctactaacaccagacatcatgtcttcaaagaccacct	120
Db	256	ACCATGAATTCAAAATCCCTGCCCTACTACACAGACATCATGTCTTCAAGAGCACCCCT	315

OY	121	gctccttcctccacaaagctcgctagaaacaagaatttbtcccaagg9gg9ccctgg9gaat	180
Db	316	GCTCTTTCACCAAAAGCTCCTGAGAACAAGATTTTGTCCAGAGGGGCCCTGGGGAAT	375
OY	181	tgttcagcaaatagtcatgatgcagagagtgltcatagtgccttgagaaaattcccacatc	240
Db	376	TGTTTCAGCAAAATGATGATCATGATGCAGAGAGTGTCATGGCTTGCAGAAATTCCCATACTCAGT	435
OY	241	gagaaagccataaacctttggcaaaaaaagctcttcttggaaaagcatcttttgaagt	300
Db	436	GAGAAAGCCATAAACCTTGGCAAGAAAAAGTCTTGTGAAAAGCATCTTTGAGTAGTG	495
OY	301	ggagaagaaattcgcagagacgcctgccaaagctctcgtctcaggtccaagagactlg	360
Db	496	GAGACGAAAGATTCGCAGAGAGCACCGCTCCCAAAGTCTCTGCTCAGGGTCCAAGAGACTTG	555
OY	361	gaataccaagaattcgcacaagcagcaagtggtctcaggtgtcttctaagctgtagc	420
Db	556	GAAATACCAGAAATTCGCACAGCCACAGCGATGATGATCTTCTTAACCGTGAGAGCACTGC	615
OY	421	ttggagcatgaagctgtggagccccaaagtatcttcattgccaaccgagtagtaaat	480
Db	616	TTGGACACATGAAGCTGTGGACCCCAAAATCATTTCCATTGCCCCAACCAGAGTAGCTGAATTT	675
OY	481	gttatcttcctggcacccaccaagcagcagcccaagcagagagagcttcaaftccaaagagat	540
Db	676	GTTTACTCCTCGGCACACCAACAAGCGAACCCAGGACGAGGGCTTCAAGTCCAAAGANATTT	735
OY	541	tttgttaactgaagggtctctctccctccoaagctccaaaggccagtgacctgtagcttcaagttct	600
Db	736	TTTTGTAACTGAGGGGTCTCTCTCTCCAGCTCCCAAGGCCACAGTGGCTGTAGCTTCAAGTTCT	795
OY	601	aagaaagatgaagaagaacaataactagccaaatgtttgaagctgctgaataattccga	660
Db	796	AAGAAAGATGAAGAGAGAACAACTACTACCAAAATTTGTTGAGCTGCTGAATATTTCAGCA	855
OY	661	gattcatgtggaagaagaag	678
Db	856	GATCAGTTGCAAGAAAG	873
RESULT	9		
AC007537		190858 bp	DNA linear PRI 09-Oct-1999
LOCUS		Homo sapiens 12P12 BAC RPCR11-267J23	(Roswell Park Cancer Institute)
DEFINITION		Human BAC library) complete sequence.	
ACCESSION		AC007537	
VERSION		AC007537.3	GI:4914348
KEYWORDS		HTG.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
		1 (bases 1 to 190858)	
		Muzny,D., Arenson,A.D., Bouck,J., Bunac,C., Chen,Z., Ding,Y.,	
		Dugan,S., Durbin,J., Forcum,J., Garcia,C., Gorrell,J.H.,	
		Gorelli,L.L., Hernandez,J., Issar,A., Jackson,L., Kneitz,S.,	
		Kondejewski,N., Lau,S., Leal,B., Lee,E., Lichtarge,O., Liu,W.,	
		Logan,O., Lu,J., Marondel,I., Martinez,C., Perschke,S., Miller,A.,	
		Montgomery,K., Oswal,G., Pampall,L.R., Parish,B.J., Perez,L.,	
		Rashtid,N.D., Rives,C., Scherer,S.E., Shen,H., Shim,C., Simon,M.,	
		Vo,Q., Williamson,A., Worley,K.C., Xiang,X.M., Yang,R., Yu,W.,	
		Zhou,X., Kucherlapati,R., Nelson,D. and Gibbs,R.A.	
TITLE		Direct Submission	
JOURNAL		Unpublished	
REFERENCE		2 (bases 1 to 190858)	
AUTHORS		Worley,K.C.	
TITLE		Direct Submission	
JOURNAL		Submitted (13-Jan-1999)	
REFERENCE		College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	
AUTHORS		3 (bases 1 to 190858)	
		Worley,K.C.	

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT

Direct Submission
Submitted (27-MAY-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 190858)
Worley,K.C.
Direct Submission
Submitted (28-MAY-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 190858)
Worley,K.C.
Direct Submission
Submitted (09-OCT-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 28, 1999 this sequence version replaced gi:4895156.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAR-REPORT-----

Contig length: Summary Statistics -----
Phrap values in estimate: 190858
Average error rate (BCM-Phrap estimate): 188824
Fraction of Phrap values less than 40 : 0.000239024
Number of consensus changing edits: 0.039746
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
16186 atattgtgaa(c)cccatctct atattgtgaa(a)cccatctct
26355 taatttga(n)acagcgcaaaa taatttga(a)acagcgcaaaa
35016 ttttttttt(n)agcggaagtc ttttttttt(g)agcggaagtc
35923 ttattatcat(n)ttttttttt ttattatcat(l)ttttttttt
49346 agcagcgcgga(n)cttcgcagtga agcagcgcgga(g)cttcgcagtga

49643 tggccaacat(n)gtgaaacc tggccaacat(a)gtgaaacc
80511 gagccgagag(n)tgacaacag gagccgagag(c)tgacaacag
91836 gaaggaaga(n)accagcttg gaaggaaga(a)accagcttg
104278 accaatlaag(n)tagtagacg accaatlaag(t)tagtagacg
114732 ctccctccct(n)cctccctccc ctccctccct(c)cctccctccc
185043 catgttgaa(n)ccgctctca catgttgaa(c)ccgctctca
189415 gatcacttga(n)cccagaagc gatcacttga(a)cccagaagc

----- Distribution of Quality < 40 Bases -----

#	bases	5	10	15	20	25	30	35	40
10001		*	*	*	*	*	*	*	*
9001		*	*	*	*	*	*	*	*
8001		*	*	*	*	*	*	*	*
7001		*	*	*	*	*	*	*	*
6001		*	*	*	*	*	*	*	*
5001		*	*	*	*	*	*	*	*
4001		*	*	*	*	*	*	*	*
3001		*	*	*	*	*	*	*	*
2001		*	*	*	*	*	*	*	*
1001		*	*	*	*	*	*	*	*
01		*	*	*	*	*	*	*	*

Version: 1.01 xfigo.
FEATURES
source Location/Qualifiers
1..190858
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12p12"
/clone="RP111-267023"
/complement(65..337)
/rpt_family="MER7C"
338..647
/rpt_family="AluY"
complement(648..961)
/rpt_family="MER7C"
1251..1423
/rpt_family="MER5B"
1598..1723
/rpt_family="MER5A"
3236..3327
/rpt_family="MIR"
complement(3840..4697)
/rpt_family="L1MA8"
5217..5484
/rpt_family="MSTC"
5485..5785
/rpt_family="AluSc"
5786..5941
/rpt_family="MSTC"
complement(5976..6074)
/rpt_family="HAL1"
6170..6191
/rpt_family="TTTTGTG)n"
complement(6284..6456)
/rpt_family="MLT1L"
complement(6577..6887)
/rpt_family="AluSx"
complement(6892..7077)
/rpt_family="HAL1"
complement(8555..8857)
/rpt_family="AluSx"
9274..9448
/rpt_family="L1MD"
9445..10362
/rpt_family="L1MEC"
10363..10656
/rpt_family="AluSx"

(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 21476: contig of 21476 bp in length
* 21477 21576: gap of unknown length
* 21577 41450: contig of 19874 bp in length
* 41451 41550: gap of unknown length
* 41551 55298: contig of 13748 bp in length
* 55299 55398: gap of unknown length
* 55399 66764: contig of 11366 bp in length
* 66765 66864: gap of unknown length
* 66865 81567: contig of 14703 bp in length
* 81568 81667: gap of unknown length
* 81668 94373: contig of 12706 bp in length
* 94374 94473: gap of unknown length
* 94474 104139: contig of 9666 bp in length
* 104140 104239: gap of unknown length
* 104240 114453: contig of 10214 bp in length
* 11454 114553: gap of unknown length
* 11455 122925: contig of 8372 bp in length
* 122926 123025: gap of unknown length
* 123026 127231: contig of 4206 bp in length
* 127232 127331: gap of unknown length
* 127332 134754: contig of 7423 bp in length
* 134755 134854: gap of unknown length
* 134855 140454: contig of 5600 bp in length
* 140455 140554: gap of unknown length
* 140555 147535: contig of 6981 bp in length
* 147536 147635: gap of unknown length
* 147636 153069: contig of 5434 bp in length
* 153070 153169: gap of unknown length
* 153170 159234: contig of 6065 bp in length
* 159235 159334: gap of unknown length
* 159335 164827: contig of 5493 bp in length
* 164828 164927: gap of unknown length
* 164928 169721: contig of 4794 bp in length
* 169722 169821: gap of unknown length
* 169822 172565: contig of 2744 bp in length
* 172566 172665: gap of unknown length
* 172666 175115: contig of 2450 bp in length
* 175116 175215: gap of unknown length
* 175216 177758: contig of 2543 bp in length
* 177759 177858: gap of unknown length
* 177859 180950: contig of 3092 bp in length
* 180951 181050: gap of unknown length
* 181051 184037: contig of 2987 bp in length
* 184038 184137: gap of unknown length
* 184138 187485: contig of 3358 bp in length
* 187486 187595: gap of unknown length
* 187596 189691: contig of 2096 bp in length
* 189692 189791: gap of unknown length
* 189792 191812: contig of 2021 bp in length
* 191813 191912: gap of unknown length
* 191913 194143: contig of 2231 bp in length.

FEATURES
source
1. 194143
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-52513"
BASE COUNT 54237 a 42367 c 41198 g 53596 t 2545 others
ORIGIN

Query Match 57.4% Score 435.8 DB 2 Length 194143;
Best Local Similarity 97.4% Pred. NO. 1.1e-12;
Matches 443; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 atgtatgacacagatggatgtatcaccctggagaagaatccctatagatgatgacctaac 60
Db 177208 ATGTATGACACAGATGGATGTATCACCCTGGAGAAGAATCCCTAGATGATGATGACTTAAC 177149
QY 61 aacatagaattcaaaatccctgactactacacacacacatcatgtctcaagaaccct 120
Db 177148 AACATAGAATTCAAAATCCCTGACTACTACACACACATCATGTCTCAAGAGCACCCCT 177089
QY 121 gctctcttccacccaagtgtctgagaagaagttgtcccaagagggtctgggaat 180
Db 177088 GCTCTCTTCTCACCAAGTGTGTGAGAACAGAGTTGTGCCAGAGGGGCTGGGGAT 177029
QY 181 ttttaagaanaatgaatcatgagcaagatgtcatgtgacctggagaattcccaatcaat 240
Db 177028 TGTTCAGCAAAATGAGTCAATGAGCAGAGGCTCATGGCTTGCAGAAATTCCTCAATCCAGT 176969
QY 241 gagaagccataaactcttgcaagaagaagctctcttggaagaatctcttggaagtg 300
Db 176968 GAGAAGCCATTAACCTTGCAAGAAAGTCTTGGAAAGCATTTCTTGAGTAGTG 176909
QY 301 gagaaggaagatctgcaagagcagctctgcaaggtctctctgaaggtcaagaagcttg 360
Db 176908 GAGAAGGAAGATTCGACAGCACGCTGCAAGGCTCTGCTCAAGGTCAAGAGACGTTG 176849
QY 361 gaataccaagatctgcaagcagcagctgtgcaagtgctctctctcaagctggagcaagctg 420
Db 176848 GAATTCACAAGATTCGACAGCACGCTGCTCAAGGCTCTTCTTAACGCTGAGACAGTGC 176789
QY 421 ttgagcatgaagctgttgaccccaagatcttc 455
Db 176788 TTGAGCATGAAGTGAAGCATGCTGAGTTCCTTTC 176754

RESULT 11
AC091816
LOCUS 200499 bp DNA linear HTG 02-Oct-2001
DEFINITION Homo sapiens chromosome 12p clone RP11-52513, WORKING DRAFT
SEQUENCE, 9 unordered pieces.
AC091816
AC091816.7 GI:15625994
VERSION
KEYWORDS HTG; PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 200499)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Banks,T., Barbara,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Boucek,J.,
Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Day-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,U.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtenarge,O., Lieu,C., Liu,J., Liu,W.,
Lounsbeg,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Maxwell,E., McLeod,M.P., Meador,M.,
Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabadi,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,

REFERENCE
AUTHORS

TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

BASE COUNT 272 a 246 c 271 g 197 t 1 others
 ORIGIN

Query Match 36.2% Score 274.4: DB 6: Length 987;
 Best Local Similarity 67.7%: Pred. No. 1.5e-72;
 Matches 469; Conservative 0; Mismatches 197; Indels 27; Gaps 5;

QY 1 atgcttagcaccagtggtgtgtgacctgtgagaagaatccccctagatgatgtgacctaac 60
 DB 1 ATGCGACGACACAGCTGTATGACTCTGGAAGACATTCCTCCGAGGATGATGATCAAC 60
 QY 61 accatagaatcaaaatccctgcctactacaccagacatcatgtcttcaagagcaccct 120
 DB 61 AGCATAGATTCAAAATCTGCTGTCTTACGCGACAGACACATCTTCAAGAACACCCCG 120
 QY 121 gctctctctacaaaactgtctgagaagaagtcttcccaaggaggcctgggaat 180
 DB 121 GCTCTCTTCTGCCCCACACTCTCCAGAACAGAGTCTGTCCAGAAAGCCCTGGGACT 180
 QY 181 tttcagaacaaatgagatcagagagagtgatgagcctgtgcaaaatcccaatccagt 240
 DB 181 TGTG---CAACTGATTCCTGGACACAGATATCATTCCTTGCAGAGGTTCCCTCCAGC 237
 QY 241 gagaagaccataaactctgtgcaagaagaagtcttcttgaagaactcttggagtagtg 300
 DB 238 GAAAGAAACATCACCTTGCGCAAGAAAGTCTTCTTGGAGAACACTCTTCAGGGTGCC 297
 QY 301 gaga---aggaagattgacagagacagcgtctgcaaggtcttctgctcaaggtaagagc 357
 DB 298 GAGAGAGAGGAGGCGCTGCGCTGCCCAAGAGATCGAGCTCAGGGTCTCTCAGGGC 357
 QY 358 ttggaataccagaattctg-----cacagccagcagtgctcgaaggtcttctt 405
 DB 358 CCGTCCCGGTAGAGGCGGACAGTGGCTTCCACACACACACAGCTGCGCCAGGCTCTAGC 417
 QY 406 aagctggaagcagtgctgtgagacatgaagctgtggaacccaagatcttccattgtgcaac 465
 DB 418 AGCTGTGAGGAGG---CCCTGAGAGTGAAGTGTGATTCCTCAAGTGTGTATGGCAAC 476
 QY 466 cgaagtagctgaatgtttatctctgcccacacacacagcagcccaagcagcagaagcttc 525
 DB 477 AGAGTGGCTGAATGTGTACTCTGCGCACACACAGATGTATCCACACAGCGAGGAGA 536
 QY 526 aagtcacaagaagatttctgaactgaaggtctctctccctcagcctcagaagcagctgct 585
 DB 537 AGCCAGCTCAAAAGAGGCTCTCGGAGATTTTGTACTTCAGGTTTGAAGAGACTTGGA 596
 QY 586 gtgactcgaagttctagaagaagaagaacaatactagccaataattgttgagctg 645
 DB 597 CTC-----TAAGAAATTAAGATGTGTGAAGCAAAATTAAGCAAAATTTGAGCTG 648
 QY 646 ctgaatatcagaagatcagttgagaagaag 678
 DB 649 CTGAATCTCGGGGATCAGTTGGGAAGAGAG 681

RESULT 13
 AC097248/c 149350 bp DNA linear HNG 20-Dec-2001
 LOCUS Rattus norvegicus clone CH230-177C11, *** SEQUENCING IN PROGRESS
 DEFINITION *** 64 unordered pieces.
 ACCESSION AC097248
 VERSION AC097248.3 GI:17973829
 KEYWORDS HNG; HNGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 149350)
 Muzny,D.M., Adams,C., Adlo-Oduola,B., Ali-Osman,F.R., Allen,C.,
 Alstbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbara,J.,

Benton,J., Biname,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Dem,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodson,A., Hogue,M., Holloway,C.,
 Hollins,B., Homs,F., Howard,S., Huber,J., Hultk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichteberg,O., Lieu,C., Liu,J., Liu,W.,
 Louisedge,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
 Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
 Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
 Ogunb,M., Okunoye,G., Oragunye,N., Oyiedo,R., Pace,A., Payton,B.,
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.,
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
 Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shoostari,N.,
 Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Taber,P., Tameris,A., Tameris,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,Y., Villalobos,J., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.

Direct Submission
 Unpublished
 2 (bases 1 to 149350)
 Worley,K.C.
 Direct Submission
 Submitted (13-Oct-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Dec 20, 2001 this sequence version replaced gi:17064657.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

----- Project Information
 Center project name: GINL
 Center clone name: CH230-177C11
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329first call to
 findphraplist

Consensus quality: 114307 bases at least Q40
 Consensus quality: 121845 bases at least Q30
 Consensus quality: 128763 bases at least Q20
 Estimated insert size: 114370; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-fp estimation
 Quality coverage: 1.5x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 64 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

Search completed: June 20, 2002, 05:11:00
Job time: 9460 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 16:39:44 ; Search time 46.53 Seconds
(without alignments)
936.916 Million cell updates/sec

Title: US-09-771-961-4
Perfect score: 1316
Sequence: 1 MCTSGCDLEETPLDDDDLTIEFKILAYTRHHVFPKSTPALFSPKLRTRTSLSGRGLGN 252

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 562222
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_protist:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriophage:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1316	100.0	252	4	Q9BZR7
2	1189.5	90.4	327	4	Q9BZR8
3	1181	89.7	276	4	Q9BZR5
4	651	49.5	328	11	Q9CPT0
5	650	49.4	328	11	Q9D3W3
6	97	7.4	450	3	013722
7	95.5	7.3	1253	11	Q9ECG8
8	95	7.2	1083	10	048839
9	93	7.1	980	4	Q9HCE4
10	91.5	7.0	960	6	018965
11	91.5	7.0	987	6	018966
12	89.5	6.8	647	16	Q9WZEO
13	89.5	6.8	696	5	Q9VK07
14	89	6.8	1252	11	Q9ECG9
15	89	6.8	1252	11	Q9JCL1
16	89	6.8	1262	5	Q9N911

17	89	6.8	1293	10	Q9AXD6	Q9AXD6 zea mays (m
18	88.5	6.7	250	11	Q9CRI8	Q9CRI8 mus musculu
19	88	6.7	638	11	Q99KW6	Q99KW6 mus musculu
20	88	6.7	4202	11	Q91V63	Q91V63 mus musculu
21	88	6.7	4731	11	Q91XP8	Q91XP8 mus musculu
22	87.5	6.6	287	3	043030	043030 schizosach
23	87.5	6.6	791	11	Q9C2X2	Q9C2X2 mus musculu
24	87.5	6.6	3115	5	Q95Y72	Q95Y72 leishmania
25	87	6.6	1660	10	Q9SIR1	Q9SIR1 arabidopsis
26	85.5	6.5	464	16	P74199	P74199 synechocyst
27	85.5	6.5	889	10	Q9MIM3	Q9MIM3 arabidopsis
28	85.5	6.5	1311	5	Q9V899	Q9V899 drosophila
29	85	6.5	1080	10	Q9V3028	Q9V3028 arabidopsis
30	84	6.4	336	10	Q93364	Q93364 brassica ol
31	84	6.4	494	15	Q9QC00	Q9QC00 human immun
32	84	6.4	724	15	083009	083009 lymphoprol
33	83.5	6.3	497	15	Q901A1	Q901A1 human immun
34	83.5	6.3	1284	5	Q23974	Q23974 drosophila
35	83	6.3	332	16	Q98Q48	Q98Q48 mycoplasma
36	83	6.3	489	11	Q9D4A0	Q9D4A0 mus musculu
37	83	6.3	1190	16	Q91XT6	Q91XT6 mus musculu
38	83	6.3	1193	16	Q9CIE1	Q9CIE1 lactococcus
39	83	6.3	1264	4	Q9HZE5	Q9HZE5 pseudomonas
40	83	6.3	1447	10	064473	064473 homo sapien
41	83	6.3	3477	11	Q99MH8	Q99MH8 mus musculu
42	83	6.3	4202	11	Q91XP9	Q91XP9 mus musculu
43	83	6.3	4731	11	Q91X00	Q91X00 mus musculu
44	83	6.3	258	10	Q9XJ56	Q9XJ56 daucus caro
45	82.5	6.3	258	10	Q9XJ56	Q9XJ56 daucus caro

ALIGNMENTS

RESULT	ID	Query Match	Score	DB 4	Length	252
1	Q9BZR7	100.0%	1316	4	252	0
AC	Q9BZR7	Best Local Similarity	100.0%	Pred. No. 2.1e-115	Indels	0
DT	01-JUN-2001 (Tremblrel. 17, Created)	Matches	252	Conservative	0	Mismatches
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)					
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)					
DE	APOPTOSIS REGULATOR BCL-2 SHORT FORM.					
GN	BCL2.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
OX	NCBI_Taxid-9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE-21264734; PubMed-11054413;					
RT	Guo B., Godzik A., Reed J.C.;					
RT	"Bcl-2, a novel pro-apoptotic member of the Bcl-2 family.";					
RL	J. Biol. Chem. 276:2780-2785(2001).					
DR	EMBL, AF281255; AAG59794.1;					
DR	SEQUENCE 252 AA; 28089 MW; 87D2E5123EFCB9E4 CRC64;					

Query Match	Score	DB 4	Length	252
Best Local Similarity	100.0%	Pred. No. 2.1e-115	Indels	0
Matches	252	Conservative	0	Mismatches
QY	1	MCTSGCDLEETPLDDDDLTIEFKILAYTRHHVFPKSTPALFSPKLRTRTSLSGRGLGN 60		
DB	1	MCTSGCDLEETPLDDDDLTIEFKILAYTRHHVFPKSTPALFSPKLRTRTSLSGRGLGN 60		
QY	61	CSANESWTEVSWPCRNSSQSEKAINLGKKSSWKAFFGVVEKEDSQSTPAKVASAGQRTL 120		
DB	61	CSANESWTEVSWPCRNSSQSEKAINLGKKSSWKAFFGVVEKEDSQSTPAKVASAGQRTL 120		
QY	121	ETQDSHQQMSRCLSNVEQCLEHFAVDPKVIISANRVAEIIYSSNPPPOATGAGFKSKEI 180		
DB	121	ETQDSHQQMSRCLSNVEQCLEHFAVDPKVIISANRVAEIIYSSNPPPOATGAGFKSKEI 180		

```

QY 181 FVTEGLSFQLOGHVPAVSSSKKDEEQLAKIVELLKTSQDOLERKDTAFIPPLVDTST 240
DB 181 FVTEGLSFQLOGHVPAVSSSKKDEEQLAKIVELLKTSQDOLERKDTAFIPPLVDTST 240
QY 241 QGFPDGLMACI 252
DB 241 QGFPDGLMACI 252

RESULT 2
Q9BZR8 PRELIMINARY: PRT: 327 AA.
AC Q9BZR8;
DB 01-JUN-2001 (Tremblrel. 17, Created)
DB 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DB 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE APOPOLOSIS REGULATOR BCL-2 LONG FORM.
GN BCLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264734; PubMed=11054413;
Guo B., Godzik A., Reed J.C.;
"Bcl-2, a novel pro-apoptotic member of the Bcl-2 family.";
RL J. Biol. Chem. 276:2780-2785(2001).
DR EMBL: AF281254; AAG59793.1;
DR InterPro: IPR002475; BCL2_family.
DR InterPro: IPR000712; Bcl-2.
DR Pfam: PF00452; Bcl-2; 1.
DR SMART: SM00337; BCL; 1.
DR PROSITE: PS00662; BCL2_FAMILY; 1.
SQ SEQUENCE 327 AA; 36598 MW; 155D43BA91BFB3EA CRC64;

Query Match 90.4%; Score 1189.5; DB 4; Length 327;
Best Local Similarity 93.5%; Pred. No. 2.2e-103;
Matches 222; Conservative 2; Mismatches 7; Indels 7; Gaps 1;

QY 1 MCSTSGCDLEIPDDDDLTIEFKILAYTRHHVKSPTALPSPKLLTRSLSQRLGN 60
DB 1 MCSTSGCDLEIPDDDDLTIEFKILAYTRHHVKSPTALPSPKLLTRSLSQRLGN 60
QY 61 CSANESWTEVSWPCRNSSSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSAGORTL 120
DB 61 CSANESWTEVSWPCRNSSSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSAGORTL 120
DB 121 EYODSHSQWRSCLSNVEQCLEHEAVDPKVISIANRVAEIVSWPPQATQAGFKSKEI 180
DB 121 EYODSHSQWRSCLSNVEQCLEHEAVDPKVISIANRVAEIVSWPPQATQAGFKSKEI 180
QY 181 FVTEGLSFQLOGHVPAVSSSKKDEEQLAKIVELLKTSQDOLERKDTAFIPPLVDTST 240
DB 181 FVTEGLSFQLOGHVPAVSSSKKDEEQLAKIVELLKTSQDOLERKDTAFIPPLVDTST 240
QY 241 QGFPDGL 248
DB 241 QGFPDGL 248

RESULT 3
Q96OR5 PRELIMINARY: PRT: 276 AA.
AC Q96OR5;
DB 01-DEC-2001 (Tremblrel. 19, Created)
DB 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DB 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE APOPOLOSIS REGULATOR BCL-2 MEDIAN FORM.
GN BCLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Montpetit A., Boily G., Sinnott D.;
RT "A detailed transcriptional map of the chromosome 12p12 tumor
suppressor locus.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY040274; AAK72109.1;
SQ SEQUENCE 276 AA; 30948 MW; 81559A7190F5598E CRC64;

Query Match 89.7%; Score 1181; DB 4; Length 276;
Best Local Similarity 99.6%; Pred. No. 1.1e-102;
Matches 226; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCSTSGCDLEIPDDDDLTIEFKILAYTRHHVKSPTALPSPKLLTRSLSQRLGN 60
DB 1 MCSTSGCDLEIPDDDDLTIEFKILAYTRHHVKSPTALPSPKLLTRSLSQRLGN 60
QY 61 CSANESWTEVSWPCRNSSSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSAGORTL 120
DB 61 CSANESWTEVSWPCRNSSSEKAINLGKKSSWKAFFGVVEKEDSOSTPAKVSAGORTL 120
DB 121 EYODSHSQWRSCLSNVEQCLEHEAVDPKVISIANRVAEIVSWPPQATQAGFKSKEI 180
DB 121 EYODSHSQWRSCLSNVEQCLEHEAVDPKVISIANRVAEIVSWPPQATQAGFKSKEI 180
QY 181 FVTEGLSFQLOGHVPAVSSSKKDEEQLAKIVELLKTSQDOLERK 227
DB 181 FVTEGLSFQLOGHVPAVSSSKKDEEQLAKIVELLKTSQDOLERK 227

RESULT 4
Q9CPT0 PRELIMINARY: PRT: 328 AA.
AC Q9CPT0;
DB 01-JUN-2001 (Tremblrel. 17, Created)
DB 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DB 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE 9030625M01RIK PROTEIN (49334055K19RIK PROTEIN).
GN 9030625M01RIK OR 49334055K19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=COLON, STOMACH, AND TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kaotaka K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochava H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK018579; BAB31290.1;
DR EMBL: AK008662; BAB25830.1;
DR EMBL: AK016670; BAB30370.1;

```


QY 160 IYVSWPPQATQAGFKSKET-FYTEGLS-----FQLOGHVPAVSSSK-----DEEQ 207
 DB 301 IPRVSPSRATLGVGASGSLTGLSRAPHYELDNGNRLSKRLHDLQIDSEK 360
 QY 208 IIAKIVELKYSGDQLEKRDPAF 230
 DB 361 IEPVELELLLEIDFEVESYVNF 383

RESULT 7

Q9E0S8 PRELIMINARY; PRT; 1253 AA.
 AC Q9E0S8;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE DM36E.
 GN NOPE OR DM36E.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6JA;
 RA Aoyama T., Nagayama S., Oka M., Kiyono T., Sasaki M.S., Nakamura T.,
 RA Toguchida J.,
 RA "Up-regulation of a ras effector and down-regulation of a cell
 RT adhesion molecule are associated with transformation of osteoblasts."
 RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB052621; BAB19279.1;
 DR HSSP; P56276; ITLK.
 DR MGD; MGI:1858497; Nope.
 DR InterPro: IPR003962; Fh11_repeat.
 DR InterPro: IPR003961; FN_111.
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003598; Ig_c2.
 DR InterPro: IPR003600; Ig_1like.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam; PR00041; fn3; 5.
 DR Pfam; PR00047; Ig; 4.
 DR PRINTS; PR00014; FNTYPE111.
 DR SMART; SM00060; FN3; 5.
 DR SMART; SM00409; IG; 4.
 DR SMART; SM00408; IGC2; 4.
 DR SMART; SM00410; IG_1like; 2.
 DR SMART; SM00410; IG_1like; 2.
 KM Immunoglobulin domain; Repeat.
 SO SEQUENCE 1253 AA; 134894 MW; 2D55B254A42B9AEE CRC64;

Query Match 7.3%; Score 95.5; DB 11; Length 1253;
 Best Local Similarity 23.2%; Pred. No. 5.4; Mismatches 82; Indels 39; Gaps 7;
 Matches 45; Conservative 26;

QY 8 DLEIPLDDDDLT-IEFKILAYYTRHHVFKSTPALF-----PKILTRTSLSORGIG 61
 DB 483 DTTELQVRLPPTVDYEFVYVAYSQLGASRTSPALVHTLDVPSAAPQLTIS----- 535
 QY 62 SANESTEVSW-PCRNSSSEKALINLKKSSWKAFFGVVEKESQSTPAKYSAQGRFL 120
 DB 536 SPNPSIRVWMLPLPSSLNSQVLF-----KYKIEYGLGKEEDQVSTEVGNETQLTL 588
 QY 121 EYQDHS-----QMSRCLSNVEQCLEHNAVDPKVISINARVAEIVYS 163
 DB 589 NSIQPKKVVYRVRISAGTACGAGVPSQWMMQHTPGVHNOSHVPAELKVKAKMESLVVS 648
 QY 164 W-PPQATQAGFK 176
 DB 649 WQPPHPTQISGK 662

RESULT 8

Q48839
 ID 048839 PRELIMINARY; PRT; 1083 AA.
 AC 048839;
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
 DE PUTATIVE UBIOUITIN CARBOXYL TERMINAL HYDROLASE.
 GN ATG32780.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RX MEDLINE=20083487; PubMed-10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Coppenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:761-768(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV_COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AC003974; AAC04485.1;
 DR MEROPS; C19; UPW; -;
 DR InterPro: IPR001394; UCH-2.
 DR InterPro: IPR001607; zf-UBP.
 DR Pfam; PF00442; UCH-1; 1.
 DR Pfam; PF00443; UCH-2; 1.
 DR Pfam; PF02148; zf-UBP; 1.
 DR SMART; SM00290; znf-UBP; 1.
 DR PROSITE; PS00973; UCH_2_2; 1.
 DR PROSITE; PS02035; UCH_2_3; 1.
 DR HydroLase.
 SO SEQUENCE 1083 AA; 120770 MW; 55B881F9971199D9 CRC64;

Query Match 7.2%; Score 95; DB 10; Length 1083;
 Best Local Similarity 21.8%; Pred. No. 4.9; Mismatches 94; Indels 90; Gaps 13;
 Matches 62; Conservative 39;

QY 8 DLEIPLDDDDLT-IEFKILAYYTRHHVFKSTPAL-----FSPKILTRTSLSORGIG 59
 DB 786 DSDRAVLDESBSVPVDRCLAQFTKHEIISDNAMCENCSKULQRLKRRKTKGLS 845
 QY 60 NCSANESTEVSW-PCRNSSSEKALINLKKSSWKAFFGVVEKESQSTPAKYSAQGRFL 101
 DB 846 NRWVWENGASSAFDECRPSLSNQCIDL---ENGXYAAPRTIKLPNCKEESATDQFVG 902
 QY 102 KEDSQSTP---AKYSAQGRFLTYO---DSHQMSRCLSNVEQCLEHNAVDPKVISIA 154
 DB 903 EENTKQAPITTSVETPLPLGGETISSQPASDNECENM-----EDLAVSEEV-IV 950
 QY 155 NRVA--EIVYSWPPQAT-----QAGGFKSKELFYTEGLSFOLOGHVPAVSSSK----- 196
 DB 951 KRDAKKVILINKAPVYLTHLKRFSODANGRVSK-----LSGHVDFQEFIDLSK 999
 QY 197 ---ASSKKDEEQLIAKIVELL-----KYSQDQLEKRD 228
 DB 1000 YMDRCSSEDEPVYRLAGVLEHLGAMSRGHVYSYINGHKKERDS 1044

RESULT 9

Q9HCE4

```

ID 09HCE4 PRELIMINARY; PRT; 980 AA.
AC 09HCE4;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE KIAA1628 PROTEIN (FRAGMENT).
GN KIAA1628.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; Pubmed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirotsawa M., Ohara O.;
RT Prediction of the coding sequences of unidentified human genes.
RT XVII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.;
EMBL: AB046848; BAB13454.1; -.
DR InterPro: IPR003962; FNIL1_repeat.
DR InterPro: IPR003961; FNIL1.
DR InterPro: IPR003598; I9_C2.
DR InterPro: IPR003600; I9_Like.
DR pfam: PF00041; fn3; 5.
DR pfam: PF00047; I9; 1.
DR PRINTS: PRO0014; FNTYPEIII.
DR SMART: SM00060; FN3; 5.
DR SMART: SM00409; I9; 1.
DR SMART: SM00408; I9C2; 1.
DR SMART: SM00410; I9_Like; 1.
KW Immunoglobulin domain; Repeat.
FT NON_TER 1
FT SEQUENCE 980 AA; 105361 MW; A0636C95D4E951E CRC64;

Query Match 7.1%; Score 93; DB 4; Length 980;
Best Local Similarity 24.4%; Pred. No. 6.6;
Matches 47; Conservative 25; Mismatches 83; Indels 38; Gaps 8;

QY 8 DLEIPLDDDDLN-IEKILAYIRHHVFKSTPALFS-----PKLLRTRSLSGRLGNC 61
DB 214 DTLELQVRDLEPNTDYEFVYVAYSQLGASRTSPALVHTLDPVPSAAPOLSL----- 266
QY 62 SANESWTEVSW-PCRNSSSEKAILGKKSSWKAFFGVYEKEDSGSTPAK-----V 112
DB 267 SPNPSDIRVAPLPPSLNSGOVY-----KKTLEYGLGKEQIPSTEVNGNETOLMLN 319
QY 113 SAAGQRTLEYQDSHS-----QQMSRLSNVEQCLEHEAVDPKVISIANRYAEIYYSW 164
DB 320 SLDPNKYVRIRISAGTAAGFAPSCWMHHRTPSMNOSHVPFAPALPKYQAKKESLVYSW 379
QY 165 -PPQATQAGGFK 176
DB 380 QPPHPPTQISGYK 392

RESULT 10
O18965 PRELIMINARY; PRT; 960 AA.
AC O18965;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE EAG CHANNEL.
GN EAG1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98190099; Pubmed=9524140;
RA Frings S., Bruell N., Dzeja C., Angele A., Hagen V., Kaupp U.B.,
RA Baumann A.;
RT "Characterization of ether-a-go-go channels present in photoreceptors
RT reveals similarity to Ix, a K current in rod inner segments.";
RL J. Gen. Physiol. 111:583-599(1998).
DR EMBL: Y13430; CAA73842.1; -.
DR InterPro: IPR000636; Cation_chan_non_11g.
DR InterPro: IPR001622; Channel_pore_K.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR001610; PAC.
DR InterPro: IPR000700; PAS-assoc_C.
DR pfam: PF00027; CNMP_binding; 1.
DR pfam: PF00520; ion_trans; 1.
DR pfam: PF00785; PAC; 1.
DR SMART: SM00100; CNMP; 1.
DR SMART: SM00086; PAC; 1.
DR SMART: SM00091; PAS; 1.
DR PROSITE: PS0042; CNMP_BINDING_3; 1.
SQ SEQUENCE 960 AA; 108092 MW; 86FCD330A3E3B08 CRC64;

Query Match 7.0%; Score 91.5; DB 6; Length 960;
Best Local Similarity 21.3%; Pred. No. 8.9;
Matches 49; Conservative 40; Mismatches 80; Indels 61; Gaps 10;

QY 15 DDDDLNTIEFKILAYIRHHVFKST-----PALFSPKLL--RTNLSLORG 57
DB 733 DDDDLVERKGSVLTESHSHGLAKASVTVRESPATPVAFPAAPAGIDHRLQAPGAE 792
QY 58 LG-----NCSANESWTEVSWPCRNSSSEKAILGKKSSWKAFFGVYEKEDSGSTPAK 111
DB 793 LEPKAGGACCAKKGVARFEDACGQAEWSK-----VSKAESMTLPBR 836
QY 112 VSAAGQRTLEYQDSHSQQMSRC--LSNVEQCLEHEAVDPKVISIANRYAEIYYSW-P 166
DB 837 TKAAGATLAKTNDSCSGITKSLRLDNVGEARSPQDRSP-----ILAEVKHSFYIP 889
QY 167 PQATQAGGSKSEIFVTEGLSFOLGHVPAVSSSKKDEEDILAKIYELL 216
DB 890 EQTLQAAVLEVKH-----ELKEDIK-ALSTKMTSIEKOLSEILRIL 929

RESULT 11
O18966 PRELIMINARY; PRT; 987 AA.
AC O18966;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE EAG CHANNEL.
GN EAG2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98190099; Pubmed=9524140;
RA Frings S., Bruell N., Dzeja C., Angele A., Hagen V., Kaupp U.B.,
RA Baumann A.;
RT "Characterization of ether-a-go-go channels present in photoreceptors
RT reveals similarity to Ix, a K current in rod inner segments.";
RL J. Gen. Physiol. 111:583-599(1998).
DR EMBL: Y13431; CAA73843.1; -.
DR InterPro: IPR000636; Cation_chan_non_11g.
DR InterPro: IPR001622; Channel_pore_K.
DR InterPro: IPR000595; CNMP_binding.
DR InterPro: IPR001610; PAC.

```



```
DR      FLybase-FBgn0032499; CG61116.  
SQ     SEQUENCE    696 AA;   80690 MW;   2E6FDA367CF8D0CD CRC64;  
  
Query Match          6.8%; Score 89.5; DB 5; Length 696;  
Best Local Similarity 21.0% Pred. No. 8.9'  
Matches 48; Conservative 29; Mismatches 77; Indels 75; Gaps 11;  
  
OY       GCDDLEIP-----LDDDDLNTIEFKILAYTRHHVFKSTPALFSPKLRTSLSGRLG 59  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB       GFNIESPHEKSELDLPD-----DVLLTYLT-HTDKASEARYTSEKLPQRHQQKAEE 80  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
OY       NGSANESWTEVSMPCRNSQSSESKEAINLGKKSKSWAFEGVEEKEDSOSTPAKSAGCORT 119  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB       ICDDEDAWRKTNNOC-----VCVKV-----WKHY--SAERRDGPDP----- 114  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
OY       LRYODSHSOQMRSCLSNVEQCLENHADVCPKVYSIANRVAIYWSM-----PPQA 169  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
115 -EVGRKHNDVFGR-----SQLRP---SRLPPELLLSWGYYFGGLPLPSLTL 159  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
OY       TTAGA-----GFKSKEIFVTGLSFLOGHVPVASSSKRDEEPQ 207  
| : | : | : | : | : | : | : | : | : | : | : | : | : |  
DB       SQCGRNCTVQLNGCEFPASPMISBOALOSQHLYHKYAEEKLEBPQ 208
```

RESULT 14

```
OYEQS9 PRELIMINARY; PRT; 1252 AA.
```

```
ID OYEQS9 AC OYEQS9 ID OYEQS9 AC OYEQS9  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2004 (TREMBLrel. 19, Last annotation update)  
DE DDM36.  
GN NOPE OR DDM36.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
NC Mammalia; Eutheria; Rodentia; Sciurogathi; Muridae; Murinae; Mus.  
OX NCBI_TaxID=10090;  
[ ]  
RA RN RPNP RP  
RC STRAIN=C57BL/CBA.  
RA Murakami H., Nakamata T., Nakayama T., Yamamoto H., Hosaka T.,  
RA Aoyama T., Nagayama S., Oka M., Kiyono T., Sasaki M.S., Nakamura T.,  
RA Toguchida J.;  
RT "Up-regulation of a ras effector and down-regulation of a cell  
adhesion molecule are associated with transformation of osteoblasts."  
Submitted (DEC-2000) to the EMBL/Genbank/DDbj databases.  
EMBL AB052620; BAB19278.1;  
HSXP F56276; ITLK.  
MGD MG1:1858497; Nope.  
DR InterPro: IPRO03962; FnIII_repeat.  
DR InterPro: IPRO03961; FN_III.  
DR InterPro: IPRO03599; Ig_  
DR InterPro: IPRO03598; Ig_c2.  
DR InterPro: IPRO03600; Ig_Like.  
DR InterPro: IPRO03006; Ig_MHC.  
DR Pfam PF00041; fn3_5.  
DR Pfam PF00047; ig_4.  
DR PRINTS PR00014; ENTYPEIIII.  
DR SMART SM0060; FN3_5.  
DR SMART SM00409; IG_4.  
DR SMART SM00408; IGC2_4.  
DR SMART SM00410; IG_Like_2.  
KW Immunoglobulin domain; Repeat.  
SEQUENCE 1252 AA; 134764 MW; BA5292393483ABV3 CRC64;
```

Query Match 6.8%; Score 89; DB 11; Length 1252;
Best Local Similarity 23.7%; Pred. No. 22;
Matches 46; Conservative 26; Mismatches 82; Indels 40; Gaps 8;

DLEEIFLDDDLNLT-IETKIATYTTHRHVFKSTPALFS-----PKLIKRTSLSORGIUGC 61
| : : : : | : || :||| | :|||| | :|||

```

Db      483 DTTELQVRDLERPNIDYEYVVAISQLGASRTSSPALVHLDDVPASAPOULTS----- 535
QY      62 SANESWTEVSW-PCRNSSOSEKAINLGGKKSSWKRAFFGVVEEKEDSOSTPAKVSAOGORTL 120
       ||| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     536 SPNPSDIRVAMLPSPSSLSNGVL-----KYKIIEYG-LGRKDQVFSTEVPGETOLT 587
QY    121 EYDSDSHS-----QQMSKCLSNVEOCLLEAVDPKVIISIANRVAELIYS 163
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     588 NSLDQPNKVYRVIRISAGTGAGYGVPSQMQRTPPGVHNOSHVPAPAEALTKRAKMESLIYS 647
QY    164 W-PPEQATQAGGFK 176
       | ||| || |||
Db     648 WOPPHPTQISGYK 661

RESULT 15
09JUL11 PRELIMINARY; PRT; 1252 AA.
ID      09JUL11
AC      09JUL11;
DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      NEIGHBOR OF PUNC ELL PROTEIN.
GN      NOPE.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OX      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=FVB;
RM      MEDLINE=20175427; PubMed=10708514;
RA      Salpaum J.M., Kappen C.;
RT      "Cloning and expression of Nope, a new mouse gene of the
RL      immunoglobulin superfamily related to guidance receptors.";
       Genomics 64:15-23(2000).
DR      EMBL; AF176694; AAF65930.1; -.
DR      HSSP; P56276; ITLK.
DR      GMD; MGJ1858497; Nope.
DR      InterPro: IPR003962; FNIII_repeat.
DR      InterPro: IPR003961; FN_III.
DR      InterPro: IPR003598; Ig_C2.
DR      InterPro: IPR003600; Ig_Like.
DR      InterPro: IPR003006; Ig_MHC.
PFam: Pfam0041; fn3; 5.
DR      Pfam: PF00047; fn3; 4.
DR      PRINTS; PR00014; ENTYPETILT.
DR      SMART; SM00060; FN3; 5.
DR      SMART; SM00408; IGc2; 3.
DR      SMART; SM00410; IG_Like; 1.
KW      Immunoglobulin domain; Repeat.
SQ      SEQUENCE 1252 AA; 134759 MW; 11948773277B76B2 CRC64;

Query Match          6.8%; Score 89; DB 11; Length 1252;
Best Local Similarity 23.7%; Pred.No. 22;
Matches 46; Conservative 26; Mismatches 82; Indels 40; Gaps
                        8;
```

Thu Jun 20 13:15:11 2002

us-09-771-961-4.rspt

Page 8

Search completed: June 19, 2002, 16:39:46
Job time: 399 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 19, 2002, 16:40:06 ; Search time 15.65 Seconds

(without alignments)
623.471 Million cell updates/sec

Title: US-09-771-961-4

Perfect score: 1316

Sequence: 1 MCTSGCDLEHPLDDDLN.....IPVDTSIQGFPPQGLMACI 252

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt.40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101.5	7.7	530	1 TDT_MOUSE	P09638 mus musculu
2	91.5	7.0	2319	1 ARA6_HUMAN	Q13023 homo sapien
3	90	6.8	433	1 KAR1_YEAST	P11927 saccharomyc
4	89.5	6.8	696	1 UVRG_DROME	Q9VX07 drosophila
5	88.5	6.7	1131	1 AC15_MOUSE	P35601 mus musculu
6	88	6.7	486	1 TECL1_YEAST	P18412 saccharomyc
7	87	6.6	520	1 TDT_BOVIN	P06526 bos taurus
8	86	6.5	639	1 GCAL_HUMAN	Q9UJY5 homo sapien
9	85.5	6.5	250	1 YC97_HUMAN	Q9Y3A5 homo sapien
10	85.5	6.5	464	1 CAP_DICDI	P54654 dictyostell
11	85	6.5	702	1 DY12_ANTCR	Q16959 antioctidari
12	85	6.5	1468	1 DRPA_YEAST	P13382 saccharomyc
13	84.5	6.4	2875	1 RPO1_TSMY1	P28976 tomato spot
14	83	6.3	1264	1 STV2_HUMAN	P26640 homo sapien
15	82.5	6.3	1468	1 N153_RAT	P49791 rattus norv
16	81	6.2	890	1 GIND_ECOLI	P27249 escherichia
17	80.5	6.1	1653	1 CLH1_YEAST	P22137 saccharomyc
18	80.5	6.1	1722	1 RBB2_HUMAN	P29375 homo sapien
19	80	6.1	177	1 R66_MICUL	P33099 micrococcus
20	80	6.1	690	1 STGB_BUCCI	P57335 buchnera ap
21	79.5	6.0	1242	1 MSH6_YEAST	Q03634 saccharomyc
22	79.5	6.0	2314	1 AKA6_RAT	Q09W07 rattus norv
23	79	6.0	445	1 FIDL_VIBPA	Q03475 vibrrio para
24	79	6.0	723	1 GGA3_HUMAN	O9N552 homo sapien
25	79	6.0	900	1 SYA_MYCG	P47534 mycoplasma
26	78.5	6.0	184	1 CAS3_MOUSE	Q02862 mus musculu
27	78.5	6.0	822	1 YF11_YEAST	P47046 saccharomyc
28	78.5	6.0	1123	1 DC11_DROME	P18169 drosophila
29	78.5	6.0	1589	1 DC13_DROME	P18171 drosophila
30	78	5.9	758	1 YMS8_YEAST	Q03825 saccharomyc
31	78	5.9	1147	1 AC15_HUMAN	P35551 homo sapien
32	77.5	5.9	777	1 BARI1_HUMAN	O99728 homo sapien
33	77.5	5.9	1263	1 STV2_MOUSE	Q92199 mus musculu

34	77.5	5.9	1428	1 YAB4_SCHPO	Q09773 schizosacch
35	77.5	5.9	1636	1 BUD3_YEAST	P25558 saccharomyc
36	77	5.9	623	1 HPC2_YEAST	Q01448 saccharomyc
37	77	5.9	630	1 ACH4_RAT	P09483 rattus norv
38	77	5.9	906	1 PRO1_SCHPO	Q12381 schizosacch
39	77	5.9	1805	1 NEST_RAT	P21263 schizosacch
40	76.5	5.8	502	1 CP12_HUMAN	P51589 homo sapien
41	76.5	5.8	518	1 TDT_MONDO	Q02789 monodelphis
42	76	5.8	486	1 YNB1_YEAST	P53742 saccharomyc
43	76	5.8	572	1 GAG_IPHA	P04023 hamster int
44	76	5.8	597	1 SYK_AQUAE	O67258 aquilex aeo
45	76	5.8	950	1 DC12_DROME	P18170 drosophila

ALIGNMENTS

RESULT 1

TDT_MOUSE STANDARD; PRT; 530 AA.

AC P09638; Q99PD0; Q99PD1;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE DNA nucleotidyltransferase (EC 2.7.7.31) (Terminal addition enzyme) (Terminal deoxynucleotidyltransferase) (TDT) (Terminal transferase).

GN DNTT OR TDT.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=86286588; PubMed=3755527;

RA Koike O., Yokota T., Kageyama T., Hirose T., Yoshida S., Arai K.;

RT "Isolation and characterization of bovine and mouse terminal deoxynucleotidyltransferase cDNAs expressible in mammalian cells."

RL Nucleic Acids Res. 14:5777-5792(1986).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/C; TISSUE=Thymus;

RX MEDLINE=93219079; PubMed=8464703;

RA Doyen N., Fanton D'Andon M., Bentolila L.A., Nguyen T.Q., Rougeon F.;

RT "Differential splicing in mouse thymus generates two forms of terminal deoxynucleotidyl transferase."

RL Nucleic Acids Res. 21:1187-1191(1993).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.

RC STRAIN=C57BL/6; TISSUE=Thymus;

RX PubMed=11136823;

RA Benedict C.L., Gillfillan S., Kearney J.F.;

RT "The long isoform of terminal deoxynucleotidyl transferase (Tdt) enters the nucleus and, rather than catalyzing N addition, modulates the catalytic activity of the short isoform."

RL J. Exp. Med. 193:89-99(2001).

RN [4]

RP PRELIMINARY CHARACTERIZATION OF ALTERNATIVE FORMS.

RX MEDLINE=96016194; PubMed=7556063;

RA Bentolila L.A., Fanton D'Andon M., Nguyen T.Q., Martinez O., Rougeon F., Doyen N.;

RT "The two isoforms of mouse terminal deoxynucleotidyl transferase differ in both the ability to add N regions and subcellular localization."

RL EMBO J. 14:4221-4229(1995).

CC -I- FUNCTION: TEMPLATE-INDEPENDENT DNA POLYMERASE WHICH CATALYZES THE RANDOM ADDITION OF DEOXYNUCLEOSIDE 5'-TRIPHOSPHATE TO THE 3'END OF A DNA INITIATOR. ONE OF THE IN-VIVO FUNCTION OF THIS ENZYME IS THE ADDITION OF NUCLEOTIDES AT THE JUNCTION (N REGION) OF REARRANGED IG HEAVY CHAIN AND T CELL RECEPTOR GENE SEGMENTS DURING THE MATURATION OF B AND T CELLS. The Tdtl isoform seems to serve to modulate the function of Tdts.

CC -I- CATALYTIC ACTIVITY: N = N diphosphate + (deoxynucleotide)(M+N).

CC -1- COFACTOR: REQUIRES MAGNESIUM.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms, rdt-L(arge)/TdtL (shown here)
CC and Tdt-S(mall)/TdtS; are produced by alternative splicing. The
CC Tdt-S form is the major form.
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-X FAMILY.
CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X04123; CA27735.1; -
DR EMBL; X68670; CA48634.1; -
DR EMBL; AF316014; AAK07884.1; -
DR EMBL; AF316015; AAK07885.1; -
DR PIR; B23595; B23595.
DR HSSP; P06766; 1BPB.
DR MGD; MGI:98659; Dntf.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR002054; DNA_POLX.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00966; DNA_PolymeraseX; 1.
DR PRINTS; PR00869; DNAPOLX.
DR SMART; SM00292; BRCT; 1.
DR SMART; SM00483; POLXC; 1.
DR PROSITE; PS50172; BRCT; 1.
DR PROSITE; PS00522; DNA_POLYMERASE_X; 1.
KW Transferrase; Nucleotidyltransferase; Terminal addition; Magnesium;
KW Nuclear protein; Alternative splicing.
FT DOMAIN 27 118
FT VARSPIC 483 502 MISSING (IN ISOFORM TDT-S).
FT CONFLICT 26 26 T -> M (IN REF. 2).
FT CONFLICT 99 99 L -> F (IN REF. 2).
FT CONFLICT 99 99 R -> G (IN REF. 1).
FT CONFLICT 183 193 Q -> K (IN REF. 1).
FT CONFLICT 297 287 E -> Q (IN REF. 1).
FT CONFLICT 309 309 D -> H (IN REF. 1).
FT CONFLICT 367 367 DRRA -> ECAC (IN REF. 1).
FT CONFLICT 441 444 RAF -> AS (IN REF. 2).
FT CONFLICT 443 445
SQ SEQUENCE 530 AA; 60331 MW; E6B109DCE39C8107 CRC64;

Query Match 7.7%; Score 101.5; DB 1; Length 530;
Best Local Similarity 23.5%; Pred. No. 0.28;
Matches 54; Conservative 41; Mismatches 92; Indels 43; Gaps 11;

8 DLEIPDDDDNTIERKILAYTRRHVFKSTPALFSKLLRTNSLSORGAGNSANDSW 67
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
170 DALDIPDIDLEKRENGSCLAFMRAVSYLKSLPPPI-----TSMKDEGI-PCIGDRVK 222
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
68 TEVSWPCNSGSSSE-KALNIGKKKKSKAF--FGVEKEDSQSPAPVAGSGRTL-EY 122
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
223 SIIEGIIEDGSSSEKAVLNDERKYSFKLFTSYGVGGLK-----TAEWFRGFTLSKI 277
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
123 QDSHSQMSRC-----LSNVEQCLEHEAVDPKVISIANRYAELIVYSW-PPQATOAGG 174
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
278 QSDSLRFTQMQAGFLYEDYLVSCVNR-----PEAEAVSMIVKKAIVFLPDALVTMTGG 333
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
334 FRRG-----KMGTHVDVFLITSPATEDEEQQLLHRYVDWKQOG 373
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :

RESULT 2
ID AKA6_HUMAN STANDARD; PRT; 2319 AA.
AC Q13023; O15028;
FT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE A-kinase anchor protein 6 (Protein kinase A anchoring protein 6)
DE (PKA6) (A-kinase anchor protein 100 kDa) (AKAP 100) (MAKAP).
GN AKAP6 OR AKAP100 OR KIAA0311.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9343692; PubMed=10413680;
RA Kapliff M.S., Shillace R.V., Westphal A.M., Scott J.D.;
RT "MAKAP: an A-kinase anchoring protein targeted to the nuclear membrane
RT of differentiated myocytes.";
RT J Cell Sci. 112:2725-2736(1999).
RN [2]
RP SEQUENCE OF 934-2319 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RT DNA Res. 4:141-150(1997).
RN [3]
RP SEQUENCE OF 1666-2319 FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=95238446; PubMed=7721854;
RA McCartney S., Little B.M., Langeberg L.K., Scott J.D.;
RT "Cloning and characterization of A-kinase anchor protein 100
RT (AKAP100). A protein that targets A-kinase to the sarcooplasmic
RT reticulum.";
RT J. Biol. Chem. 270:9327-9333(1995).
CC -I- FUNCTION: BINDS TO TYPE II REGULATORY SUBUNITS OF PROTEIN KINASE A
CC AND ANCHORS/TARGETS THEM TO THE NUCLEAR MEMBRANE OR SARCOPLASMIC
CC RETICULUM. MAY ACT AS AN ADAPTER FOR ASSEMBLING MULTIPROTEIN
CC COMPLEXES.
CC -I- SUBUNIT: INTERACTS WITH RII SUBUNIT OF PKA, PHOSPHATASE 2B
CC (CALCINEURIN) AND AKAP79.
CC -I- SUBCELLULAR LOCATION: SARCOPLASMIC RETICULUM AND NUCLEAR MEMBRANE
CC IN HEART MUSCLE. PARTICIPATION OF MULTIPLE TARGETING SIGNALS ALLOW
CC CORRECT INTRACELLULAR TARGETING. THESE MAY BE REPEATED MOTIFS RICH
CC IN BASIC AND HYDROPHOBIC AMINO ACIDS, PALMITOYLATED/MYRISTOYLATED
CC MOTIFS OR ALTERNATIVELY SPICE TARGETING SEQUENCES.
CC -I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN CARDIAC AND SKELETAL
CC MUSCLE, FOLLOWED BY BRAIN.
CC -I- DOMAIN: RII-ALPHA BINDING SITE, PREDICED TO FORM AN AMPHIPATHIC
CC HELIX. COULD PARTICIPATE IN PROTEIN-PROTEIN INTERACTIONS WITH A
CC COMPLEMENTARY SURFACE ON THE R-SUBUNIT DIMER.
CC -I- SIMILARITY: CONTAINS 2 SPECTRIN REPEATS.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: U17195; AAA92354.2; -
DR EMBL: AB002309; BAA20770.1; -
DR MIM: 604691; -
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00435; spectrin; 2.
DR SMART: SM00150; SPEC; 2.
KW Repeat.
FT REPEAT 762 848 SPECTRIN 1.
FT REPEAT 1036 1150 SPECTRIN 2.
FT DOMAIN 1560 1701 SER-RICH.
FT DOMAIN 2063 2076 PKA-RII SUBUNIT BINDING DOMAIN.
FT CONFLICT 974 974 C -> W (IN REF. 2).

FT CONFLICT 1492 1492 V -> A (IN REF. 2).
SEQUENCE 2319 AA; 256663 MM; 3D3037AE4DFCCEE CRC64;

Query Match
Best Local Similarity 21.9%; Pred. No. 13; Length 2319;

Matches 51; Conservative 38; Mismatches 71; Indels 73; Gaps 11;

QY 77 SSGSEKAINLGKK-----KSSWKAFFGVKEKEDSOTPAKVSAGORTLEYOD 124
Best Local Similarity 21.9%; Pred. No. 13; Length 2319;

Db 699 SSSSDIASLGEISIESGPLDILSEESSMPLAGKKYADEKSERASSS-----EKNE 751

QY 125 SHSOQSRCSLNWEOCLEH-----EAVDPKVISIANRVAEIVY-----SWPPQAT 170

Db 752 SHSATKSAALQKLMQDIOHODNYEAIWEKIEGFVNKLDEFIOMLNAMEETTEMTPPKAE 811

QY 171 QAGGFKSKIEFVTEGSLRQL-----OGH--VYPAASSK-----KDEEQLAK 211

Db 812 -----MDLKLTYLTHLSKLVNDSHCALKEAVEEGHOLLELASHKAGIKDLMRTIASQ 867

QY 212 IVELLKXSGDLERK-----DTAFIPIPLVDTSIQ-----GFPQDGLMAC 251

Db 868 WKELQR-----QIKRQHSWILRALDITIKAEILATDVSEDEEGTGFKAEOQLC 916

RESULT 3
KARL YEAST
ID KARL YEAST STANDARD; PRT: 433 AA.

AC P11977;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Cell division control protein KARL.

KARL OR YNL188W OR N1611.

OS Saccharomyces cerevisiae (Baker's yeast).

OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.

OX NCBI_TaxID=4932;

RA SEQUENCE FROM N.A.

RP MEDLINE=87159524; PubMed=3030557;

RA Rose M.D., Fink G.R.;

RT "KARL, a gene required for function of both intranuclear and

extranuclear microtubules in yeast.";

RL Cell 48:1047-1060(1987).

SEQUENCE FROM N.A.

RP Obermaier B., Piravandi E., Rinke M., Domdey H.;

Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

FUNCTION.

RA MEDLINE=94350199; PubMed=8070654;

RA Vallen E.A., Ho W., Winey M., Rose M.D.;

"Genetic interactions between CDC31 and KARL, two genes required for

duplication of the microtubule organizing center in Saccharomyces

cerevisiae.";

RL Genetics 137:407-422(1994).

FUNCTION.

RA MEDLINE=95181541; PubMed=7876310;

RA Spang A., Courtney I., Grein K., Matzner M., Schiebel E.;

"The Cdc31p-binding protein Karip is a component of the half bridge

of the yeast spindle pole body.";

RL J. Cell Biol. 128:863-877(1995).

FUNCTION.

RA KARL IS REQUIRED FOR FUNCTION OF BOTH INTRANUCLEAR AND

EXTRANUCLEAR MICROTUBULES. KARL HELPS LOCALIZE CDC31 TO THE

SPINDLE POLE BODY (SPB). CDC31 THEN INITIATES SPB DUPLICATION VIA

INTERACTION WITH A DOWNSTREAM EFFECTOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M15683; AAA34716.1; -

DR EMBL; Z71464; CAH96082.1; -

DR PIR; A26455; A26455.

DR SGD; S0005132; KARL.

KW Cell division: Microtubules; Mitosis.

FT CONFLICT 199 199 A -> V (IN REF. 1).

SEQUENCE 433 AA; 50653 MM; D65331B3B2B06B4 CRC64;

Query Match
Best Local Similarity 21.4%; Pred. No. 2.1; Length 433;

Matches 55; Conservative 43; Mismatches 107; Indels 52; Gaps 10;

QY 8 DLEIPLDDDL-NTIEKILAYTRHHVFKSPALFSPKILTRSLSGRL--GNCSA 63

Db 131 DEDNKKIDEDNIENELQF-----TPKIEASTLRSLGLGRVYLVNTRNPKS 176

QY 64 NSWTVEVSNPCRSOSSEKAINLGKKSSWKAFFGVYKEKEDSOTPAKVSAGORTLEY 122

Db 177 KESHIKVKPDIINNKSORSKSAALRKQKGLPLPLPINSFNSDPTLQKKEVEVTDV 236

QY 123 -----ODSHSOQSRCSLN-----VEOCL-----EHAVDPKVISIAN-----RVAEIV 161

Db 237 IOKKRELISSKMHRLLFHDKMKWEKLESUREYRKMPPGTGVSSSEQDSQNFISTPT 296

QY 162 YSW-----PPQATGAGF-----KSKIEFVTEGSLFQGHVPAVSSSKRDEEQL 209

Db 297 KSYSLDEKPLPLSAMNNFNDVTDNKEKETNNILKFGQORPLQLQSEIEHMTKKL 356

QY 210 AKIVELKXSGDLERK 226

Db 357 DTIELLKDDTDSKERK 373

RESULT 4
UVRG_DROME
ID UVRG_DROME STANDARD; PRT: 696 AA.

AC Q9YK07;

DT 01-MAR-2002 (Rel. 41, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE UVRAG-like protein.

GN CG6116.

OS Drosophila melanogaster (Fruit fly).

OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

CC Ephydridae; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RA SEQUENCE FROM N.A.

RP STRAIN=BERKELEY;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,

Amaratides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

Abri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,

Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,

Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

Dubin K.T., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,

Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,

DR InterPro: IPR000862; RRC.
 DR Pfam: PF00004; AAA; 1.
 DR Pfam: PF00533; BRCT; 1.
 DR SMART: SM00292; BRCT; 1.
 DR PROSITE: PS00172; BRCT; 1.
 DR DNA replication; ATP-binding; Transcription regulation; DNA-binding;
 KW Activator; Nuclear protein; Zinc-finger.
 FT DOMAIN 399 477
 FT NP_BIND 635 642
 FT ZN_FING 734 751
 FT DOMAIN 1104 1108
 FT DOMAIN 354 528
 FT CONFLICT 66 66
 FT CONFLICT 187 187
 FT CONFLICT 254 254
 FT CONFLICT 559 559
 FT CONFLICT 614 614
 FT CONFLICT 945 945
 FT CONFLICT 1071 1071
 FT CONFLICT 1104 1104
 SO SEQUENCE 1131 AA; 125984 MW; A6FAF970A7FE9E94 CRC64;

Query Match 6.7%; Score 88.5; DB 1; Length 1131;
 Best Local Similarity 23.7%; Pred. No. 9.7;
 Matches 58; Conservative 34; Mismatches 92; Indels 61; Gaps 11;

QY 48 LRRSLSORGL-----GNCNSNEMSTEVSW-----PCNNSQSSSEKAINIGKR 89
 DB 462 LGRIIDEDGLDIRTPMPGRSKYEMAEAMKREKSKLETPQKNDGKRK-ISPARK 520
 QY 90 KSSWKAFFGVKEKEDSOSTPAKVSAGORTLEYODSHSOQMSRCLSNVEQCL-----EHEA 145
 DB 521 EESKSKCKLTLLKNSPKMAVKKKASTCPRGIDYVKETHGNR-----SSNKEKCLLWVDKYKR 576
 QY 146 VDKPVT-----SIANVAEIVYSW-----PPQATGAGGKSEIETEGLSIQ--- 189
 DB 577 ASLKNIIGQGDQSCANLLRLWLRNMHKSPEEKRAAFKGLASKD---DSSFFKAL 632
 QY 190 LGHVPAVSSSKKDEEEDILAKIVELLYSGDQLERKDT-----AFIPILVDSIOG 242
 DB 633 LSGPPVGYKT-----TASLVCOELGYIVELNADSTRKNSLKAVAESINNTSIKG 685
 QY 243 FPQDG 247
 DB 686 FYTSG 690
 RESULT 6
 TECL YEAST STANDARD; PRT: 486 AA.
 ID TECL YEAST
 AC P18412;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE TY transcription activator TECL.
 GN TECL OR ROCI OR YBR083W OR YBR0750.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RX MEDLINE=90287143; PubMed=2192259;
 RA Laloux I., Dubois E., Dewechin M., Jacobs E.;
 RT "TECL, a gene involved in the activation of TYI and TYI-mediated gene
 RT expression in Saccharomyces cerevisiae: cloning and molecular
 RT analysis.";
 RL Mol. Cell. Biol. 10:3541-3550(1990).
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;

RA Andre B., Cziepluch C., Hein C., Jauniaux J.C., Urrestarazu A.,
 RA Vissers S.;
 RL Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP TEA DOMAIN.
 RX MEDLINE=91300541; PubMed=2070413;
 RA Buerklin T.R.;
 RT "The TEA domain: a novel, highly conserved DNA-binding motif";
 RL Cell 66:11-12(1991).
 CC -1 FUNCTION: TECL IS INVOLVED IN THE ACTIVATION OF TYI AND TYI-
 CC MEDIATED GENE EXPRESSION. IT IS NOT INVOLVED IN MATING OR
 CC SPOGULATION PROCESSES.
 CC -1 SUBCELLULAR LOCATION: Nuclear.
 CC -1 SIMILARITY: CONTAINS 1 'TEA' DNA-BINDING DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: M32797; AAA35141.1; -
 DR EMBL: Z35952; CAA85028.1; -
 DR PIR: A35667; A35667.
 DR TRNSPAC: T01084; -
 DR SGD: S0000287; TECL.
 DR InterPro: IPR000818; TEA.
 DR Pfam: PF01283; TEA; 1.
 DR PRINTS: PR00065; TEADOMAIN.
 DR SMART: SM0065; TEA; 1.
 DR PROSITE: PS00554; TEA.DOMAIN; 1.
 KW Transcription regulation; Trans-acting factor; Activator;
 KW DNA-binding; Nuclear protein; TEA-DOMAIN.
 FT DNA_BIND 127 192
 SO SEQUENCE 486 AA; 55157 MW; F247016D3E75C454 CRC64;

Query Match 6.7%; Score 88; DB 1; Length 486;
 Best Local Similarity 22.6%; Pred. No. 3.6;
 Matches 58; Conservative 36; Mismatches 95; Indels 68; Gaps 12;

QY 3 STSGCDLEIPDDDDINTTEFKILAYTRHNYK-----STPALFSPK 46
 DB 48 STACVKNKA---EDNINILD-----THQGEIVNTGSGAKSDLKSAKATFDK 96
 QY 47 LIRTR-----SLSORGLGNCNSNEMSTEVSWPCRNSSQSEK-----AINIGKKSSMK- 94
 DB 97 QKKNEVPNITSVSNYPFGQSESTSTTE-SWTIGCDKSKSEKVEAFLEALRLIMKNGTKI 155
 QY 95 ----AFGVKEKEDSOSTPAKVSAGORTLEYODSHSOQMSRCLSN-VEQCLEHRAVPPK 149
 DB 156 KIRNANFG---RNELISLYTKHKTNEPRTKQJSSHIQWVKTKIQNKIDSLTSSKEE 212
 QY 150 VISIANRVA-----ETVSWPPQATGAGGKSEIETV---TEGLSPQLQ 191
 DB 213 LHLLEHGAHQTTENSINLEFDIEEILDSL--PSVDSGSLTPKNIIVYNNSSGSLSVSK 270
 QY 192 GHVPAVSSSKKDEEEOI 208
 DB 271 LTPITASNEKKIENFI 287
 RESULT 7
 TDT BOVIN STANDARD; PRT: 520 AA.
 ID TDT BOVIN
 AC P06526;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA nucleotidylexotransferase (EC 2.7.7.31) (terminal addition enzyme)
 DE (Terminal deoxynucleotidyltransferase) (TDT) (Terminal transferase).
 DE

[illegible]

Db	234	C1EE1----	IEDGESESEVNAVUNDERGQSKLTTSV-----	GVG-LTKISEK	276
OY	67	WTEVSWPCNSOSSSEKAINLGKKSSMKAFEGVVEKEDSOSTPAKVASAGORFLEYODSH	126		
Db	277	WFRMGFRSLSKIMSDTKLTKFTKMQ-----KAGF-----	LYEDLV	311	
OY	127	SQOWMSCLSNVEQCLEHEAVDPVISAIVANVAELVYSM--PPQATQAGF-KSKEIFVTE	184		
Db	312	S-----CVTRAEAE-----EAV-----GVLVKAERVAWFLDPAFYTMGFFRRGKTI----	351		
OY	185	GLSFOLGCH-----VPVASSSKDEEEOILAKIVELLKSG-----DQLEKRDTPA-IPIL	235		
Db	352	-----GHNDVDFLITSPCSAEDEEDQLPVRINLMEKKGILLLYDYDVESFFENKRLPSRQ	404		
OY	236	VDT	238		
Db	405	VDT	407		
RESULT 8					
GGAL	HUMAN	STANDARD:	PRT:	639	AA.
ID	GGAL_HUMAN				
AC	O9UJY5; O9UGW0; O9UGW1; O9UGW2; O9BWM94;				
DT	16-OCT-2001 (Rel. 40; Created)				
DT	16-OCT-2001 (Rel. 40; Last sequence update)				
DT	16-OCT-2001 (Rel. 40; Last annotation update)				
DE	ADP-ribosylation factor binding protein GGAL (Golgi-1-localized, gamma				
DE	ear-containing, ARF-binding protein 1) (Gamma-adaptin related protein				
GN	GGAL.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=20214818; PubMed=107449927;				
RA	Boman A.L., Zhang C.-J., Zhu X., Kahn R.A.;				
RT	"A family of ADP-ribosylation factor effectors that can alter				
RT	transport through the trans-Golgi.";				
RL	Mol. Biol. Cell 11:1241-1255(2000).				
RN	[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RC	TISSUE=Heart;				
RX	MEDLINE=20211638; PubMed=10747089;				
RA	Dell'Angelica E.C., Puertollano R., Mullins C., Aguilar R.C.,				
RA	Vargas J.D., Hartnell L.M., Bonifacio J.S.;				
RT	"Gags: A family of ADP ribosylation factor-binding proteins related to				
RT	adaptors and associated with the Golgi Complex.";				
RL	J. Cell Biol. 149:81-94(2000).				
RN	[3]				
RP	SEQUENCE FROM N.A. (ISOFORM 1).				
RX	MEDLINE=20211637; PubMed=107447088;				
RA	Hirst J., Lui W.-Y., Bright N.A., Totty N., Seaman M.N.J.,				
RA	Robinson M.S.;				
RT	"A family of proteins with gamma-adaptin and VHS domains that				
RT	facilitate trafficking between the trans-Golgi network and the				
RT	vacuole/lysosome.";				
RL	J. Cell Biol. 149:67-80(2000).				
RN	[4]				
RP	PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).				
RA	Matthews L.;				
RL	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	SEQUENCE OF 31-639 FROM N.A. (ISOFORM 1).				
RC	TISSUE=Testis;				
RA	Wambutt R., Heubner D., Mewes H.-W., Gassenhuber J., Wiemann S.;				
RL	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.				
RN	[6]				
RP	SEQUENCE OF 272-369 FROM N.A.				
RC	TISSUE=Neuroblastoma;				
RA	Strausberg R.;				
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.				

RA Saiz J.E., Buitrago M.J., Soler A., del Rey F., Revuelta J.L.;
 RT "The sequence of a 21.3 kb DNA fragment from the left arm of yeast
 RT chromosome XIV reveals IED4, MET4, POL1, RAS2, and six new open
 RT reading frames.";
 RL Yeast 12:403-409(1996).
 RN (3)
 RP SEQUENCE OF 1-26 FROM N.A.
 RC STRAIN=5288C;
 RX MEDLINE=93188702; PubMed=8446029;
 RA Mountain H.A., Bystrom A.S., Koroch C.;
 RT "The general amino acid control regulates MET4, which encodes a
 RT methionine pathway-specific transcriptional activator of
 RT Saccharomyces cerevisiae.";
 RL Mol. Microbiol. 7:215-228(1993).
 CC -1- FUNCTION: POLYMERASE ALPHA IN A COMPLEX WITH DNA PRIMAASE IS A
 CC REPLICATIVE POLYMERASE.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + (DNA)(N).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
 CC ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
 CC DIFFERENT REACTIONS OF DNA SYNTHESIS.
 CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J03268; AAA3488.1; -;
 CC EMBL: 250161; CAA90524.1; -;
 CC EMBL: 271378; CAA95978.1; -;
 CC EMBL: 212126; CAA78111.1; -;
 CC PIR: A30211; A30211.
 CC SGD: S0005046; POL1.
 CC InterPro: IPR002064; DNA_pol_B.
 CC Pfam: PF00135; DNA_pol_B_1.
 CC Pfam: PF03104; DNA_pol_B_exo; 1.
 CC PRINTS: PRO0106; DNAPOLB.
 CC SMART: SM00486; POLBC; 1.
 CC PROSITE: PS00116; DNA_POLYMERASE_B; 1.
 CC K1M Transferrase: DNA-directed DNA polymerase; DNA replication;
 CC KM DNA-binding: Nuclear protein.
 CC DNA_BIND 1246 1381 POTENTIAL.
 CC VARIANT 493 493 G -> R (IN TEMPERATURE SENSITIVE MUTANT).
 CC CONFLICT 759 760 MI -> IV (IN REF. 1).
 CC SEQUENCE 1468 AA; 166808 MW; 50C9032DBE95B5AE CRC64;
 SQ

Query Match 6.5%; Score 85; DB 1; Length 1468;
 Best Local Similarity 25.9%; Pred. No. 27;
 Matches 51; Conservative 29; Mismatches 93; Indels 24; Gaps 8;

QY 65 ESWTEVSWPCRNSSOSSEKAINLGGKKSSWKAFFGVYKEDSOSTPAKVSQOGORTLEYOD 124
 DB 74 EDWREVD-----NSSSDEDTGNLASKDSKRRK--NIKREKHDTTDMRTKSHSTLLAHA 127
 QY 125 SHSQQMSRCLSNVEQCL--EHEAVDPKVIASIANRYAEIVYSPPOATQAGFKS--KE 179
 DB 128 KKSQSKIPDNFDIDLEEFESGEVEKPNILLPSKLRNLNSPTSE-----FKSSIKR 181
 QY 180 IFVTEGLSFG--LGGHVPVASSSKKDEEOILAKIYEL----LKYSGDQLERKDTAIP 233
 DB 182 VNGDESSHDAGISKRVKIDPDSSTDKYLESSPLKIQSKRLRYANDYDLDLV-ENS 240
 QY 234 PLVDTSIGFPODGLMA 250
 DB 241 PVVATKRGVNQLDITLLA 257

RESULT 13
 ID RRP1_TSWV1
 AC P28976;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
 GN L.
 OS Tomato spotted wilt virus (strain Brazilian Br-01) (TSWV).
 OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tospovirus.
 OX NCBI_TaxID=36413;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91374019; PubMed=1895058;
 RA de Haan P., Kormelink R., de Oliveira Resende R., van Poelwijk F.,
 RA Peters D., Goldbach R.;
 RT "Tomato spotted wilt virus L RNA encodes a putative RNA polymerase.";
 RL J. Gen. Virol. 72:2207-2216(1991).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA](N).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D10066; BAA00955.1; -;
 CC PIR: J01335; RRVUTW.
 CC Transferrase: Nucleotidyltransferase; RNA-directed RNA polymerase.
 KM SEQUENCE 2875 AA; 331498 MW; 574B3B55C0766FB CRC64;
 SQ

Query Match 6.4%; Score 84.5; DB 1; Length 2875;
 Best Local Similarity 22.0%; Pred. No. 71;
 Matches 40; Conservative 40; Mismatches 73; Indels 29; Gaps 7;

QY 80 SEKATLGGKKSSWKAFF---GVYKEDSOSTPAKVSQOGORTLEYODSHSQMSRCLIS 135
 DB 1760 NENALN---KMSYKDFSKLPLMKNEIDLYSKRNKLIKIDDAVLE---EDELTKKIAS 1812
 QY 136 INEQCLEHEAV--DPVVISIAN-----RYAEIYSPPPQATQAGFKSKEIFTEGLS 187
 DB 1813 SLEMESVHDIMKNPETILTAPLNDRDFLSQLFMYTSPSKRNQLSNOSTEKALADRVLR 1872
 QY 188 FLOGHVPVASSSKKDEEOILAKIYELKYSGD-----OLEKRDYAF-IPPIYVD 237
 DB 1873 SKARFTVNISSTYKMYEENMEKKIEMLKFDLDYSYCFKCVNLVINKVNSMLIPILD 1932
 QY 238 TS 239
 DB 1933 SA 1934

RESULT 14
 ID STV2_HUMAN
 AC P26640; Q9U0M2;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Valyl-tRNA synthetase 2 (EC 6.1.1.9) (Valine--tRNA ligase 2) (VALRS
 DE 2) (G7a).
 GN VARS2 OR G7A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-91378943; PubMed-1898367;
 RA Hsieh H.-L., Campbell R.D.;
 RT "Evidence that gene Gta in the human major histocompatibility complex
 RT encodes valyl-tRNA synthetase.";
 RL Biochem. J. 278:809-816(1991).
 RN [2]
 RP ERRATUM.
 RA Hsieh S.-L., Campbell R.D.;
 RL Biochem. J. 281:879-879(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
 RA Lasky S., Hood L.;
 RT "Sequence of the human major histocompatibility complex class III
 RT region";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Shihna S., Tamiya G., Oka A., Inoko H.;
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 201-1263 FROM N.A.
 RA MEDLINE-93154582; PubMed-8428657;
 RA Vilalta A., Donovan D., Wood L., Vogeli G., Yang D.C.H.;
 RT "Cloning, sequencing and expression of a cDNA encoding mammalian
 RT valyl-tRNA synthetase.";
 RL Gene 123:181-186(1993).
 CC -1- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
 CC + L-valyl-tRNA(Val).
 CC -1- ENZYME REGULATION: CAN BE REGULATED BY PROTEIN KINASE C-DEPENDENT
 CC PHOSPHORYLATION.
 CC -1- SUBUNIT: FORMS HIGH-MOLECULAR-MASS AGGREGATES WITH ELONGATION
 CC FACTOR 1.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC -1- SIMILARITY: THE N-TERMINAL DOMAIN IS SIMILAR TO ELONGATION
 CC FACTOR 1-GAMMA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X59303; CAA41990.1; -;
 DR EMBL: AF134726; AAD21819.1; -;
 DR EMBL: AF000503; BAB63303.1; -;
 DR EMBL: M98326; AAA81332.1; -;
 DR PIR: S17675; S17675.
 DR HSSP: P96142; IGAX.
 DR MIM: 604137; -;
 DR InterPro: IPR004046; GST_C.
 DR InterPro: IPR004045; GST_N.
 DR InterPro: IPR002300; tRNA-synt_1a.
 DR InterPro: IPR001412; tRNA-synt_1i.
 DR InterPro: IPR002303; tRNA-synt_val.
 DR Pfam: PF00043; GST_C; 1.
 DR Pfam: PF02798; GST_N; 1.
 DR Pfam: PF00133; tRNA-synt_1; 1.
 DR PRINTS: PRO0986; TRNASYNTHAL.
 DR PROSITE: PS00178; AA.TRNA.LIGASE.I; 1.
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Phosphorylation.
 FT DOMAIN 1 7200 EF-1-GAMMA LIKE.
 FT SITE 344 354 "HIGH" REGION.
 FT SITE 862 864 "KMSK" REGION.
 FT BINDING 865 865 "ATP (BY SIMILARITY)".
 FT CONFLICT 51 51 P -> S (IN REF. 1).
 FT CONFLICT 331 331 A -> G (IN REF. 5).
 FT CONFLICT 590 590 V -> G (IN REF. 1).

FT CONFLICT 620 640 ALINVPPELGLRFEARRAV -> GPHCASAFPGPAOYL
 FT CONFLICT 792 792 RPGRKC (IN REF. 1).
 FT CONFLICT 1064 1064 S -> F (IN REF. 1).
 FT CONFLICT 1169 1169 M -> I (IN REF. 5).
 FT CONFLICT 1169 1169 MISSING (IN REF. 5).
 SQ SEQUENCE 1264 AA; 140475 MW; 95CCDBB3AB148AD CRC64;

 Query Match 6.3%; Score 83; DB 1; Length 1264;
 Best Local Similarity 22.1%; Pred No. 33;
 Matches 49; Conservative 35; Mismatches 86; Indels 52; Gaps 11;

 QY 31 TRHHVKSPTALPSP-----KLRTRSLSGRGICGNSANSMTEVSPCR 75
 DB TRHHMRGFTLWNPDCDHAGIATGVVVEKKLMREGLSHHQG---REAFLEGVWK 420
 QY 76 NSQSEKAINLKKRSS--W-KAFEGVEKESQSPRAKVSAGQRLFLQDSHQMSR 132
 DB 421 BEKGRIYHQLKSLSSLDWDRACFTMDPKLSAAVTEAFVRLH-EEGIIYRSTFLVWMS 479
 QY 133 CLSNVEGCEHNAVDPKVISINRVAEIYSWPPQATQAGFRSK-EIFVTEGLSPQL 191
 DB 480 TLNSAISDIE--VDKKELT-----GRTLSVP-----GIKEKVEGVLSFAIKVQ 523
 QY 192 GHVPYASSSKKDEBQILAKIVELKYSQD---QLERKDTAF 230
 DB 524 G-----SDSDEVVAVATRIETMLGDAVAVHPKQDTRY 556

 RESULT 15
 N153_RAT STANDARD; PRT; 1468 AA.
 AC PA9791;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Nuclear pore complex protein Nup153 (Nucleoporin NUP153) (153 kDa
 DE nucleoporin).
 GN NUP153.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 602-613; 622-645 AND 971-993.
 RC STRAIN=BUFAFO; TISSUE=Liver;
 RX MEDLINE-93137325; PubMed-8422679;
 RA Sukegawa J., Blobel G.;
 RT "A nuclear pore complex protein that contains zinc finger motifs,
 RT binds DNA, and faces the nucleoplasm.";
 RL Cell 72:29-38(1993).
 CC -1- FUNCTION: POSSIBLE DNA-BINDING SUBUNIT OF THE NUCLEAR PORE
 CC COMPLEX (NPC). THE REPEAT-CONTAINING DOMAIN MAY BE INVOLVED IN
 CC ANCHORING COMPONENTS OF THE PORE COMPLEX TO THE PORE MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX. LOCATED TO THE
 CC TERMINAL RING STRUCTURE OF THE NUCLEOPORIN CAGE.
 CC -1- DOMAIN: CONTAINS X-F-X-F-G REPEATS.
 CC -1- PFM: PHOSPHORYLATED IN INTERPHASE. HYPERPHOSPHORYLATED DURING
 CC MITOSIS. MAY PLAY A ROLE IN THE REVERSIBLE DISASSEMBLY OF THE
 CC NUCLEAR PORE COMPLEX DURING MITOSIS.
 CC -1- SIMILARITY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATS
 CC SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST
 CC NUP1, NSP1, POM 121 AND MAMMALIAN P62.
 CC -1- SIMILARITY: CONTAINS 4 RANBP2-TYPE ZINC FINGERS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL; L06821; -; NOT_ANNOTATED_CDS.
DR InterPro; IPR001876; znf-RanBP.
DR Pfam; PF00641; znf-RanBP; 4.
DR SMART; SM00547; znf_RBP; 4.
DR PROSITE; PS01358; ZF_RANBP2_1; 4.
DR PROSITE; PS0199; ZF_RANBP2_2; 4.
DR Nucleic acid binding protein; Repeat; Zinc-finger; DNA-binding;
KW Phosphorylation.
FT DOMAIN 4 15 GLY-RICH.
FT ZN_FING 442 446 POLY-GLY.
FT ZN_FING 657 687 RANBP2-TYPE 1.
FT ZN_FING 721 750 RANBP2-TYPE 2.
FT ZN_FING 789 818 RANBP2-TYPE 3.
FT ZN_FING 846 875 RANBP2-TYPE 4.
SQ SEQUENCE 1468 AA; 152823 MW; C3DFD9697C556A7C CRC64;

Query Match 6.38; Score 82.5; DB 1; Length 1468;
Best Local Similarity 23.38; Pred. No. 45; Mismatches 103; Indels 39; Gaps 8;
Matches 54; Conservative 36; Mismatches 103; Indels 39; Gaps 8;
QY 32 RHVFKSTPALFSP--KLTTRSLQRLGNGSAN-----ESWTEVSWPCRNSSSEKAI 84
DB 275 RQNKVRSTP-YQAPVRRQMKAKQLNAQSYGVTSIARRILQSLKMSPLADAKRIPSAV 333
QY 85 N-----LGKKSSWKAFFGVVEKEDSQ-----STPAKVSAGORTLEYODSHSOOWS 131
DB 334 SSPINSPLDKRGIDSTYQAKKEKVDQPPYQRLMTPKPKVSIATNRTYFKPSLTPSGD 393
QY 132 RCLSNVEQCLEHEAVDPKVISIANRVAELIVSWPPQATQAGGFKSEIFVTEGLSPQLQ 191
DB 394 LRKTNORIDKKNSTVDEKNISRONREQESGFSYPNFSIPANGLSSG---VGGGGGKMRR 450
QY 192 GHVPAVSSSKKDEFEQLAKTYELKYSGDQLERKDTAFIPPLVDTSTIOGF 243
DB 451 ERTTHEVAVSKPSEEEV---EVPLLPQ-----ISLPISSSSLPTF 487

Search completed: June 19, 2002, 16:40:08
Job time: 381 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 19, 2002, 16:33:40 ; Search time 31.42 Seconds
(without alignments)
770.672 Million cell updates/sec

Title: US-09-771-961-4
Perfect score: 1316
Sequence: 1 MCGTSCGDEIEIPLDDDLN.....IPLVDTISIGCFPPDGLMACI 252

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	99.5	7.6	529	B23595	DNA nucleotidylexo
2	97	7.4	450	T37702	hypothetical prote
3	95	7.2	1083	T00790	ubiquitin-specific
4	90	6.8	433	S63143	cell division cont
5	89.5	6.8	647	G72346	hypothetical prote
6	88.5	6.7	1131	A49393	activator 1 large
7	88	6.7	486	A35667	Ty transcription a
8	87.5	6.6	287	T40340	hypothetical prote
9	87	6.6	547	A23595	DNA nucleotidylexo
10	87	6.6	1660	A84647	hypothetical prote
11	85.5	6.5	464	S75831	hypothetical prote
12	85.5	6.5	889	T47311	hypothetical prote
13	85	6.5	1080	T00587	probable ubiquitin
14	85	6.5	1468	S58250	DNA-directed DNA p
15	84.5	6.4	2875	RRVUTW	genome polypeptid
16	84	6.4	336	T14437	inner envelope mem
17	83.5	6.3	519	AD0663	hypothetical prote
18	83.5	6.3	1284	T13168	probable potassium
19	83	6.3	332	H90576	lipoprotein d [imp
20	83	6.3	1190	P86677	pyruvate-flavodoxi
21	83	6.3	1193	F83264	hypothetical prote
22	83	6.3	1265	S17675	valine--trna ligas
23	83	6.3	1447	T00530	hypothetical prote
24	82.5	6.3	869	G82499	clpb protein VCA01
25	82.5	6.3	1148	A49651	replication factor
26	82.5	6.3	1468	A44345	nucleoporin - rat
27	82.5	6.3	2241	T20971	hypothetical prote
28	82.5	6.3	2261	T20978	hypothetical prote
29	82	6.2	1121	I38127	phosphoprotein pho

30	82	6.2	1123	2	A72311	conserved hypothet
31	81.5	6.2	372	2	G96818	FKB20.3 [imported]
32	81.5	6.2	583	2	T16007	hypothetical prote
33	81.5	6.2	663	2	T49685	RNA polymerase II
34	81	6.2	715	2	D85087	hypothetical prote
35	81	6.2	890	2	G64740	[protein-Pil] urid
36	81	6.2	1647	2	T49412	hypothetical prote
37	80.5	6.1	333	2	T23324	hypothetical prote
38	80.5	6.1	851	2	AG2469	hypothetical prote
39	80.5	6.1	1653	2	A36349	clathrin heavy cha
40	80.5	6.1	1722	1	I78879	retinoblastoma bin
41	80	6.1	178	2	S29886	ribosomal protein
42	80	6.1	427	2	T26103	hypothetical prote
43	80	6.1	496	2	G64527	hypothetical prote
44	80	6.1	690	2	E84945	glycine--trna liga
45	80	6.1	1075	2	B96508	hypothetical prote

ALIGNMENTS

RESULT 1
B23595
DNA nucleotidylexotransferase (EC 2.7.7.31) - mouse
N:Alternate names: terminal addition enzyme; terminal deoxyribonucleotidyltransferase
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1988 #sequence_revision 02-Jun-1988 #text_change 18-Jun-1999
C:Accession: B23595; S30235
R:Koizumi, O.; Yokota, T.; Kageyama, T.; Hirose, T.; Yoshida, S.; Arai, K.
A:Title: Isolation and characterization of bovine and mouse terminal deoxynucleotidyl
Nucleic Acids Res. 14, 5777-5792, 1986
A:Reference number: A93633; MID:8626588
A:Accession: B23595
A:Molecule type: mRNA
A:Residues: 1-529 <KOI>
A:Cross-references: GB:X04123; NID:954765; PIDN:CAA27735.1; PID:954766
R:Doyen, N.; d'Andon, M.F.; Bentolila, L.A.; Nguyen, O.T.; Rougeon, F.
A:Title: Differential splicing in mouse thymus generates two forms of terminal deoxyn
Nucleic Acids Res. 21, 1187-1191, 1993
A:Reference number: S30235; MID:93219079
A:Accession: S30235
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-25, 'M', 27-98, 'F', 100-192, 'R', 194-286, 'O', 288-308, 'E', 310-366, 'D', 368-44
C:Superfamily: DNA nucleotidylexotransferase
C:Keywords: alternative splicing; magnesium; nucleotidyltransferase; nucleus

Query Match 7.6%; Score 99.5; DB 2; Length 529;
Best Local Similarity 23.5%; Pred. No. 0.81;
Matches 54; Conservative 40; Mismatches 93; Indels 43; Gaps 11;

QY 8 DDEIPLDDDLNTIEFKILAYTRHHVFKSTPALFSPKLLTRTSLSORGJGNCANSMSW 67
DB 170 DALDILAEDELRENGSCIAFMGASSVLKSLPPI-----TSMKIDREGI-PCIGDKV 222
QY 68 TEVSWPCRNSSQSE-KAINDGKKSSKAF---FGYEKEDSOSTPAVSAOGQTL-EX 122
DB 223 STIEGIEDEGESEAKAVLNDERKSFLETSYGVGLK-----TAEMFRGFTLSKI 277
QY 123 QSHSQQWSSRC-----LSNVEQCLEHEAVDPKVISIANRVAETIVSW-PPQATQGG 174
DB 278 QSDKSLRFTRKMKAGFLTYEDLVSCVR---PDAQVSMIVKEAVVTFLPDALVTMGG 333
QY 175 FSKSEIFVEGLSFQLOGH---VPVASSSKKDEEOTILAKIVELTKYSG 220
DB 334 FRNG-----KMTGHDVDFLTSPKATEDEEQLLHKVTHFWKQGG 373

RESULT 2
T37702
hypothetical protein SPAC15A10.02 - fission yeast (Schizosaccharomyces pombe)

submitted to the Protein Sequence Database, August 1994
A:Reference number: S45893

A:Accession: S45950

A:Molecule type: DNA

A:Residues: 1-486 <AND>

A:Cross-references: EMBL:Z35952; NID:9536345; PIDN:CAAB5028.1; PID:9536346; GSPDB:GN0000

C:Gene: SCD:TEC1; MIPS:YBR083w

A:Cross-references: SGD:S0000287; MIPS:YBR083w

A:Map position: 2R

C:Superfamily: Ty transcription activator TEC1; TEA DNA-binding domain homology

C:Keywords: DNA binding; nucleus; transcription regulation

E:123-193/Domain: TEA DNA-binding domain homology <TEA>

Query Match 6.7%; Score 88; DB 1; Length 486;
Best Local Similarity 22.6%; Pred. No. 7.6;
Matches 58; Conservative 36; Mismatches 95; Indels 68; Gaps 12;

3 STSGCDLEETPLDDDDLTIEFKILAYTRHHVFK-----STPALFSRK 46

48 STAAACVKNKA---EDNINLID-----THPQELVMTGLGAKSODLKSPSKATFTDK 96

47 LKTR-----SLSGRLGNCANESWTEVSPCRNSQSEK-----AINLKKKSKWK 94

97 QRKNVNPISVSNVFPQSSSTSTTE-SMTIGCDKSEKEAEALRLIMKNGTTKI 155

95 ----AFGEVKEKEDOSTPAKVASAOGORTLEYQDSHOSQMSCLSN-VQCCLEHVAIDPK 149

156 KINNAFG---RNEILILYIKHKTNERTTKOISHIOWWKTTONKIKDLSLTSKEKE 212

150 VISIANRVA-----EIYVSWPPQATQAGFKSEKIFV---TEGLSFQDQ 191

213 LHLILHGAQTENSMLFVDIFEIIDL-S-PYSVDSGSLPKMLIYVNSNSGLSVHAK 270

192 GAVPVASSKKDEEQT 208

271 LRLPITASNKKIENFI 287

RESULT 8

T40340
Hypothetical protein SPBC3B9.01 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T40340

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z21922

A:Accession: T40340

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Cross-references: EMBL:AL022070; PIDN:CAAL17781.1; GSPDB:GN00067; SPDB:SPBC3B9.01

A:Experimental source: strain 972n-; cosmid C3B9

C:Genetics:

A:Gene: SPDB:SPBC3B9.01

A:Map position: 2

Query Match 6.6%; Score 87.5; DB 2; Length 287;
Best Local Similarity 21.2%; Pred. No. 4.1;
Matches 52; Conservative 41; Mismatches 113; Indels 39; Gaps 10;

12 IPDDDDLTIEFKILAYTRH--HFKSTPALSPKLLRTSLTSORGNCNSANSWTE 69

56 VPLDQKE---IAFDNLEMLVEHDNANNLVPLQLMPRLIKOLESEPESTLRLAA--WT- 108

70 VSMPCNNSQSSKAINLKKKSSWKAFFGVVEKEDOSTPAKV-----SAOGORTLTX 122

109 IATVAGNNPNSQAL---IENDGKILFEGALKEDSDTKNNVLAITSKLINENGIAL 165

123 QDSHOSQMSRCLSNVEQCLEHVAIDPKVIST--ANRVAEIYVSWPPQATQAGFKSEK 180

Db 166 LDKIPNSWEMLEIIE--LKHSVMTKRVIFEFYALLIQE-----DKSKOI 208

Qy 181 FVTBGLSFQDQGHVPAVSSSKDEEQIILAKIYELLKYSQDOLERKDPFIPLYPTSI 240

Db 209 ILQKHEHQPPEKVOFSLSHSVDEDCVTKSLHTLTLFQKNKSVANTNELLSLVQPKS 268

Qy 241 QGFPQ 245

Db 269 E-FPE 272

RESULT 9

DNA nucleotidyltransferase (EC 2.7.7.31) long form - bovine

N:Alternate names: terminal addition enzyme; terminal deoxyribonucleotidyltransferase

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A23595; A60467; A05052; A29948; A32198; PC2127; PC2128; I45884

R:Koiwai, O.; Yokota, T.; Kageyama, T.; Hirose, T.; Yoshida, S.; Arai, K.

Nucleic Acids Res. 14, 5777-5792, 1986

A:Title: Isolation and characterization of bovine and mouse terminal deoxynucleotidyl

A:Reference number: A93633; MUID:86286588

A:Accession: A23595

A:Molecule type: mRNA

A:Residues: 1-302,321-510,520-547 <KOI>

A:Cross-references: GB:X04122; NID:9767; PIDN:CAA27734.1; PID:9768

A:Note: It is uncertain whether Met-1 or Met-12 is the initiator

R:Evans, R.K.; Besch, C.M.; Coleman, M.S.

Biochemistry 28, 713-720, 1989

A:Title: Photoaffinity labeling of terminal deoxynucleotidyl transferase. Identificat

A:Reference number: A60467; MUID:89229071

A:Accession: A60467

A:Molecule type: protein

A:Residues: 143-179,'X',181-187,'XXXXXX',194-199,'XX',202-230,'XXX',234-303,322-363,'

R:Peterson, R.C.; Cheung, L.C.; Mattaliano, R.J.; Chang, L.M.S.; Bollum, F.J.

Proc. Natl. Acad. Sci. U.S.A. 81, 4363-4367, 1984

A:Title: Molecular cloning of human terminal deoxynucleotidyltransferase.

A:Reference number: A93995; MUID:84272638

A:Accession: A05052

A:Molecule type: protein

A:Residues: 335-348,'X',350-355,366-377,'XX',380-381,'X',383-386,521-526,'X',528-533,

R:Pandey, V.; Modak, M.J.

J. Biol. Chem. 263, 3744-3751, 1988

A:Title: Biochemistry of terminal deoxynucleotidyltransferase. Affinity labeling and

A:Reference number: A29948; MUID:88153669

A:Accession: A29948

A:Molecule type: protein

A:Residues: 221-226,'X',228-231,'X',235-249 <PAN>

R:Pandey, V.N.; Modak, M.J.

J. Biol. Chem. 264, 867-871, 1989

A:Title: Biochemistry of terminal deoxynucleotidyl transferase. Identification and unl

A:Reference number: A32198; MUID:89093157

A:Accession: A32198

A:Molecule type: protein

A:Residues: 221-226,'X',228-231,'X',235-249 <PA2>

R:Takahara, K.; Hayashi, N.; Fujita-Sagawa, K.; Morishita, T.; Hashimoto, Y.; Noda, A

Biosci. Biotechnol. Biochem. 58, 786-787, 1994

A:Title: Alternative splicing of bovine terminal deoxynucleotidyl transferase cDNA.

A:Reference number: PC2127; MUID:94264411

A:Accession: PC2127

A:Molecule type: DNA

A:Residues: 303-320 <TAK>

A:Cross-references: GB:D17451

A:Accession: PC2128

A:Molecule type: DNA

A:Residues: 511-519 <TA2>

A:Cross-references: GB:D17452

R:Koiwai, O.; Kaneda, T.; Morishita, R.

Biochem. Biophys. Res. Commun. 144, 185-190, 1987

A:Title: Analysis of human terminal deoxynucleotidyl transferase cDNA expressible in

A:Reference number: I45884; MUID:87213162

A:Accession: I45884

A:Map position: 14L
 C:Superfamily: herpesvirus DNA-directed DNA polymerase
 C:Keywords: nucleotidyltransferase

Query Match 6.5%; Score 85; DB 2; Length 1468;

Best Local Similarity 25.9%; Pred. No. 62;
 Matches 51; Conservative 29; Mismatches 93; Indels 24; Gaps 8;

QY 65 ESMTEVSWPCRNSSQSEKAINLGKKSSWKAFFGVKEKEDSOSTPAKVSAGQRTLEYODSHSOQWSRCLS 124
 DB 74 EDWREYD---NSSSEDDGNLASKDSKRKK--NIREKHQITDMLKTOHSHKSTLLAHA 127
 QY 125 SHSOQWSRCLSVEOCL---EHEAVDPKVISIANRVAEIVYSWPPQATQAGFKS--KE 179
 DB 128 KKSOKSPIDNFDLGEFESEGEVEKPNILPSKIRENLNSPTSE-----FKSSIKR 181
 QY 180 IFVTEGLSPQ--LQGHVPVASSSKKDEEQLAKIVEL---LKYSGDQLERKDTAFIPI 233
 DB 182 VNGNDESSHDAGISKVKIDPDSDTKYLESSPLKLSRKLRKYANDVDLDDV-ENS 240
 QY 234 PLVDTSTIQGFPODGLMA 250
 DB 241 PVVATKRQNVLDQDTLLA 257

RESULT 15

RRVUTW

genome polyprotein - tomato spotted wilt virus (strain BR-01)

N:Alternate names: L protein

N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)

C:Species: tomato spotted wilt virus

C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998

C:Accession: J01335

R:de Haan, P.; Kormelink, R.; de Oliveira Resende, R.; van Poelwijk, F.; Peters, D.; Gol
 J. Gen. Virol. 71, 2207-2216, 1991

A:Title: Tomato spotted wilt virus L RNA encodes a putative RNA polymerase.

A:Reference number: J01335

A:Accession: J01335

A:Molecule type: genomic RNA

A:Residues: 1-2875 <DEH>

C:Genetics:

A:Map position: segment L

C:Superfamily: bunyavirus RNA-directed RNA polymerase

C:Keywords: nucleotidyltransferase

Query Match 6.4%; Score 84.5; DB 1; Length 2875;

Best Local Similarity 22.0%; Pred. No. 17e+02;
 Matches 40; Conservative 40; Mismatches 73; Indels 29; Gaps 7;

QY 80 SEKAINLGKKSSWKAFF---GVVEKEDSOSTPAKVSAGQRTLEYODSHSOQWSRCLS 135
 DB 1760 NENALN---KMSYKDFSKLYPNLKKNEDLYKSTNKLKIDEDAVLE---EDELKTKIAS 1812
 QY 136 NVEQCLEHEAY--DPKVYSIAN-----RVAEIVYSWPPQATQAGFKSKEIFVTEGLS 187
 DB 1813 SLEMSYVHDIMKNPFTLLPDLNDRDQLLSQLFMYTSPSKRMQLNSOSTEKALDRLRLR 1872
 QY 188 FOLQGHVPVASSSKKDEEQLAKIVELKYSGD-----QLERKDTAF--IPIPLVD 237
 DB 1873 SKARTEFVNISSTVATYENNEKKILEMKLPDLDSYCSFKTCVNLVIKIDVNFSLIPILD 1932
 QY 238 TS 239
 DB 1933 SA 1934

Search completed: June 19, 2002, 16:33:44
 Job time: 77 sec

